

The Necessary Trialectics: Reference, Inference, and Preference as the Functional Ground for Meaningful Propositional Attitudes

Muhammad Fajar Ismail

vadjhars@gmail.com

This work by Muhammad Fajar Ismail is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

To view a copy of this license, visit: <https://creativecommons.org/licenses/by-nc-nd/4.0/>

Abstract

This paper investigates the necessary functional architecture underlying complex cognitive states such as belief, desire, hope, and fear, specifically those characterized as Meaningful Propositional Attitudes About Something Not Itself (MPAASNI). MPAASNI are defined as states possessing directed propositional content, susceptibility to systematic processing, and an internal stance conferring functional significance within a system. We identify three fundamental functional capacities required for such states: Reference (R), providing directed content; Inference (I), providing systematic processing and relation of contents; and Preference (P), providing the internal stance, weighting, and functional significance. The core contribution of this paper is the articulation and defense of the Trialectics, defined as the specific functional mode of operation characterized by systematic and constitutive mutual conditioning among R, I, and P. Through rigorous, descriptive conceptual analysis, we argue that this Trialectic functioning—this deep functional integration where each capacity's state and operation conditions and is conditioned by the others—is not merely one possible arrangement but constitutes a necessary condition for the possibility of any system exhibiting MPAASNI. The necessity is demonstrated by showing that the functional integration inherently required by the definition of MPAASNI logically entails the mutual conditioning described by the Trialectics. This framework, derived non-empirically, applies to attitudes directed towards both tangible and non-tangible domains and posits a fundamental structure necessary for intentionality, rationality, and other sophisticated cognitive phenomena dependent on meaningful propositional attitudes.

1. Introduction

1.1. The Problem of Complex Cognitive States

Our engagement with the world, both as individuals and potentially as collective systems, is characterized by a rich tapestry of internal states that mediate our understanding, guide our actions, and shape our experience. Among these internal states, a particular class stands out due to its complexity, its functional significance, and the philosophical puzzles it presents: the class encompassing states commonly referred to as beliefs, desires, hopes, fears, intentions, doubts, and the like. These are often grouped under the technical philosophical term *propositional attitudes*, a category signifying mental states directed towards propositions – the abstract, truth-evaluable contents often associated with declarative sentences or the objects of thought. The central focus of this paper begins with acknowledging the philosophical challenge inherent in providing a comprehensive account of these specific, complex cognitive states.

What precisely is it to *believe* that a certain state of affairs obtains, to *desire* a particular outcome, or to *intend* to perform an action? These states are not merely fleeting sensations or simple reflexes; they possess an intricate structure and play a fundamental role in the cognitive economy of any system exhibiting them, bridging the gap between internal processing and interaction with an environment, whether that environment be physical, social, abstract, or conceptual.

The complexity of these states, which we will preliminarily designate as *Complex Cognitive States (CCS)* for initial discussion before refining the terminology, stems from several intertwined features that demand philosophical explanation. Firstly, such states exhibit *intentionality*, a philosophical term often traced back through Franz Brentano's work (*Psychology from an Empirical Standpoint*, 1874), signifying their inherent "aboutness" or "directedness" towards something. When one believes, desires, or fears, the state is not merely an undirected internal occurrence; it is *about* something specific – a possible state of affairs, an object, a property, another individual, or even another cognitive state. This directedness, this capacity for internal states to relate to things beyond themselves, is a core feature requiring explanation. How is it possible for a state within a system, whether that system is biological, artificial, or perhaps even collective, to be *about* the weather outside, a historical event, a mathematical theorem, or the potential actions of another agent? This problem of intentionality, concerning the nature and possibility of mental directedness, is a foundational challenge in the philosophy of mind and language.

Secondly, these complex cognitive states typically involve *propositional content*. While not all mental states might be propositional (e.g., raw sensations), states like belief and desire are standardly understood as taking propositions as their objects. To believe is to believe *that* something is the case; to desire is to desire *that* something be the case. The content of the attitude is structured in a way that can be evaluated for truth (in the case of belief) or for realization (in the case of desire). This structure allows these states to enter into logical relationships with one another – beliefs can be consistent or inconsistent, one belief might entail another, a belief and a desire might provide the basis for an intention. Accounting for these states requires explaining how systems can form, entertain, and manipulate these structured, truth-evaluable contents. This connects the problem to questions in logic, philosophy of language concerning the nature of propositions, and philosophy of mind concerning how such abstract contents can be realized or represented within a physical or functional system. Accessible introductions to these interconnected issues can be found in works like Jaegwon Kim's *Philosophy of Mind* (Third Edition, 2010) or William Lycan's *Philosophy of Language: A Contemporary Introduction* (Third Edition, 2018).

Thirdly, these states possess *functional significance* or *meaningfulness* within the system's overall operation. Beliefs, desires, intentions, and fears are not epiphenomenal decorations on a cognitive architecture; they are typically understood to play causal roles in guiding behavior, facilitating reasoning, enabling planning, and contributing to the system's adaptation to its environment. Beliefs inform actions based on how the world is taken to be; desires motivate actions aimed at changing the world to align with preferred states; intentions coordinate future behavior. This functional role implies that these states are integrated within a larger system, interacting with perception, memory, reasoning mechanisms, and action-guiding processes. Understanding these states requires not only analyzing their content and directedness but also characterizing their functional profile – what they *do* within the system, how they arise, how they interact with other states, and how they contribute to the system's overall behavior or internal economy. This functional perspective is central to many

contemporary approaches in philosophy of mind, particularly functionalism (see Ned Block's edited volume *Readings in Philosophy of Psychology, Vol. 1*, 1980 for classic papers).

The philosophical problem, therefore, is to provide an account that can adequately explain how systems can possess states exhibiting this specific confluence of features: directedness towards something (intentionality), structured propositional content, and functional significance within the system's economy. Explaining any one of these features presents significant challenges; explaining their integration within single states like belief or desire is the core task addressed by theories of propositional attitudes. The need for such an explanation arises from the centrality of these states to our understanding of ourselves and other complex agents, and from the apparent difficulty in reconciling these features with a purely physical or mechanistic view of the world. How do mere physical processes give rise to states that are *about* things, structured like propositions, and functionally significant in guiding behavior towards goals? This constitutes the fundamental problem space from which this paper departs.

1.2. Standard Approaches and Potential Gaps

Philosophical inquiry has generated several influential approaches attempting to address the problem of complex cognitive states, particularly propositional attitudes. These approaches offer distinct frameworks for understanding the nature of intentionality, content, and functional role, often emphasizing different aspects of the phenomenon. A brief, descriptive overview of some standard approaches helps situate the framework to be developed in this paper.

One dominant family of approaches falls under the umbrella of *Representational Theories of Mind (RTM)*, often associated with functionalism. RTM posits that propositional attitudes are relations between an agent and internal mental representations that possess semantic properties (meaning, reference, truth conditions). Holding a belief, on this view, involves having a representation with a specific propositional content stored or active within the system in a way characteristic of belief (e.g., available to guide inference and action). Key proponents like Jerry Fodor (*The Language of Thought*, 1975) argued for a "language of thought" hypothesis, suggesting that these internal representations have a combinatorial syntax and semantics analogous to natural language, allowing for the construction of complex propositional contents and enabling systematic reasoning processes that operate on the syntactic structure of these representations. Intentionality is explained by the semantic properties of these internal representations – they are *about* things because they refer to them or represent them. Functional significance is explained by the causal roles these representations play within the system's computational architecture, interacting with inputs, outputs, and other internal states according to functional rules. RTM attempts to provide a naturalistic account compatible with physicalism by grounding mental states in structured internal representations with causal powers.

Another influential approach, often associated with Daniel Dennett (*The Intentional Stance*, 1987), focuses on the *interpretive* nature of attributing propositional attitudes. According to the intentional stance, we attribute beliefs, desires, and other attitudes to a system (whether a person, animal, or artifact) when treating it *as if* it were a rational agent acting on the basis of such states allows us to successfully predict and explain its behavior. On this view, propositional attitudes might not correspond directly to discrete internal representations in the way RTM suggests, but rather represent patterns in behavior that are best understood by

adopting this interpretive strategy. The reality of beliefs and desires is tied to the predictive success of the intentional stance. Intentionality is thus a feature of the interpretive model we apply, and functional significance lies in the model's utility for prediction and explanation. This approach emphasizes the holistic and normative aspects of attitude attribution, focusing on rationality as a key component.

A third strand involves *Phenomenological Approaches*, stemming from Husserl and developed by subsequent thinkers. These approaches often prioritize the first-person, subjective experience of intentionality and consciousness. They analyze the structure of experience itself, describing how consciousness is always directed towards objects (noesis-noema structure) and how meaning and content are constituted through acts of consciousness. While not always focused on providing a third-person functional or representational account in the manner of RTM or the intentional stance, phenomenology offers rich descriptions of the "lived experience" of having beliefs, desires, and perceptions, emphasizing the way the world is presented to the subject through these intentional states. Work like Jean-Paul Sartre's analysis of belief and bad faith (*Being and Nothingness*, 1943) explores the complex relationship between belief, self-awareness, and freedom from a first-person perspective.

These standard approaches provide valuable insights into different facets of complex cognitive states. RTM offers a framework for understanding content representation and processing mechanisms. Interpretive approaches highlight the role of rationality and prediction in our understanding and attribution of attitudes. Phenomenology details the subjective structure of intentional experience. However, from the perspective motivating this paper, there may be a relative lack of emphasis on the *necessary functional integration* and *mutual conditioning* among the core capacities underlying these states. While RTM might posit distinct modules for representation and computation, and interpretive approaches focus on holistic patterns, the specific way in which the capacity for directed content (Reference), the capacity for systematic processing (Inference), and the capacity for internal orientation/weighting (Preference) must functionally interlock and mutually depend on each other to *constitute* a meaningful propositional attitude might remain underspecified or implicitly assumed rather than explicitly analyzed as a core requirement.

Standard functionalist accounts, for example, might describe belief as a state with certain causal relations to inputs, outputs, and other states, involving both representational content and inferential processing. They might also incorporate motivational states like desire. Yet, the focus is often on defining the *role* of the state rather than detailing the *necessary interdependencies* of the underlying capacities (R, I, P) that *enable* a state to fulfill that role in the specific way characteristic of a meaningful attitude. The potential gap, therefore, lies not necessarily in denying the importance of representation, reasoning, or motivation/value, but in lacking a dedicated framework that analyzes their functional *interdependence* and *mutual conditioning* as the specific, necessary basis for the possibility of meaningful propositional attitudes about something not itself. This paper proposes to fill this conceptual space by focusing squarely on this interrelation, termed the Trialectics.

1.3. Introducing the Triadic System (R, I, P)

To investigate this interrelation, this paper focuses on three fundamental functional capacities hypothesized as necessary components within any system capable of exhibiting meaningful propositional attitudes about something not itself. These capacities are conceptualized at a functional level, describing *what* the system must be able to do, rather than

specifying *how* these functions are physically or algorithmically implemented. These three core capacities constitute the Triadic System under investigation:

1. **Reference (R):** This is defined functionally as the capacity of the system to establish and utilize connections between its internal states (e.g., concepts, symbols, representations) and specific entities, properties, relations, or states of affairs (whether internal or external, tangible or non-tangible). The function of Reference is to provide *directed content*, enabling the system's internal states to be *about* something determinate. This capacity grounds the intentionality of propositional attitudes. It encompasses the ability to form singular terms, general concepts, relational predicates, and ultimately structured propositional contents that purport to represent the world or other domains. The specifics of how these connections are established (causally, descriptively, etc.) are less critical at this stage than the functional role of providing directed content.
2. **Inference (I):** This is defined functionally as the capacity of the system to process and relate contents, deriving new contents from existing ones according to systematic patterns or rules. The function of Inference is to provide *systematic structure and processing*, enabling the system to move beyond isolated contents, establish logical or evidential links between propositions, evaluate consistency, draw conclusions, and engage in reasoning. This capacity encompasses deductive, inductive, and abductive patterns, providing the framework for manipulating and evaluating propositional content. It grounds the systematicity and productivity often associated with thought and reasoning involving propositional attitudes.
3. **Preference (P):** This is defined functionally as the capacity of the system to exhibit an *internal stance, weighting, orientation, or selection* regarding contents, states of affairs, or processing pathways. The function of Preference is to provide the specific *mode* of a propositional attitude (distinguishing belief from desire, hope from fear, etc.) and to imbue the attitude with *functional significance* within the system's economy by assigning weights, priorities, or value-orientations that influence subsequent processing and potential action. This capacity encompasses epistemic criteria (valuing truth, simplicity), cognitive tendencies (inherent processing biases), and practical/affective orientations (desiring outcomes, fearing possibilities). It introduces the element of internal valuation or disposition towards the contents handled by Reference and processed by Inference.

These three functional capacities – R for directed content, I for systematic processing, P for internal stance/weighting – are posited as the fundamental building blocks whose specific interrelation constitutes complex cognitive states like propositional attitudes. They are presented here descriptively, based on an analysis of the apparent functional requirements for states like belief and desire as discussed in standard philosophical analyses.

1.4. The Central Thesis

Having introduced the core problem (explaining complex cognitive states/MPAASNI) and the fundamental functional capacities (R, I, P) involved, this paper advances a specific thesis concerning their relationship. The thesis arises from a detailed conceptual analysis (to be presented in subsequent sections) of the necessary conditions for the possibility of MPAASNI, based on the described functional roles of R, I, and P. The central thesis is:

"The Trialectic Functioning of a System is a Necessary Condition for the Possibility of that System Having Meaningful Propositional Attitudes About Something Not Itself."

The Trialectics, as will be elaborated in detail, refers to the specific mode of functional interrelation characterized by the mutual conditioning among the capacities for Reference, Inference, and Preference. It describes a state where the operation of each capacity is dependent upon, constrained by, and influenced by the concurrent operation of the other two, forming an integrated functional architecture. The thesis claims that this specific, integrated mode of functioning is not merely incidental but is fundamentally required for a system to constitute states that possess the combined features of directed propositional content, systematic processing, and internal functional significance that characterize meaningful propositional attitudes about something not itself.

1.5. Methodology

The derivation and defense of this thesis rely solely on a specific methodology: descriptive conceptual analysis. This approach involves:

- Defining the key concepts (Reference, Inference, Preference, Trialectics, MPAASNI) based on their functional roles and interrelations.
- Analyzing the necessary conditions for the possibility of the target phenomenon (MPAASNI) based on these functional definitions.
- Examining the logical entailments of these definitions and the described interrelations.
- Maintaining strict neutrality regarding normative evaluation (e.g., not claiming the Trialectics is 'good' or 'rational', merely describing its function) and avoiding direct appeal to empirical findings from psychology, neuroscience, or artificial intelligence as primary evidence for the conceptual claims about necessary conditions (though the concepts themselves are informed by the phenomena studied by these fields).

The aim is to demonstrate the necessity of the Trialectics *within the conceptual framework established by the paper's definitions and descriptive analyses*. The strength of the argument lies in its internal coherence and the logical connection between the description of the Trialectics and the description of the necessary functional requirements for MPAASNI. This method is suited for exploring the fundamental functional architecture required for complex cognitive phenomena at an abstract level, investigating the conditions of possibility prior to, or independent of, specific empirical implementations.

1.6. Roadmap

The remainder of this paper will unfold as follows: Section 2 will provide a more detailed functional definition and elaboration of the three components of the Triadic System: Reference, Inference, and Preference. Section 3 will articulate the concept of the Trialectics, meticulously describing the specific modes of mutual conditioning and interdependency among R, I, and P. Section 4 will offer a precise definition of the target phenomenon, Meaningful Propositional Attitudes About Something Not Itself (MPAASNI), detailing its constitutive functional requirements. Section 5 will present the core argument demonstrating the necessity of the Trialectics for the possibility of MPAASNI, drawing directly from the conceptual analysis in the preceding sections. Section 6 will explore the scope of this necessity, discuss how particular attitude types manifest the Trialectic architecture, and identify other complex phenomena rendered impossible without it according to the framework. Section 7 may briefly

consider potential conceptual objections and responses within the framework. Finally, Section 8 will conclude by summarizing the findings and restating the central thesis, highlighting its potential philosophical significance for understanding the foundations of complex cognition.

2. Defining the Components of the Triadic System

Having established in the Introduction the philosophical challenge posed by complex cognitive states, particularly Meaningful Propositional Attitudes About Something Not Itself (MPAASNI), and having proposed that the key to understanding their possibility lies in the specific interrelation of three fundamental functional capacities, this section undertakes the task of providing detailed, analytical, and functionally grounded definitions for these three components: Reference (R), Inference (I), and Preference (P). The aim here is not to offer exhaustive reviews of the vast philosophical literatures surrounding each concept, nor to commit to specific theories about their underlying mechanisms or implementations. Instead, the goal is to articulate, with maximal clarity and minimal arbitrariness, the precise functional role that each capacity *must* fulfill within any system capable of exhibiting MPAASNI, based on the constitutive requirements identified in the preliminary analysis. These functional definitions serve as the necessary groundwork for understanding the Triadics – the specific mode of mutual conditioning among R, I, and P – which will be argued in subsequent sections to be the necessary condition for the possibility of MPAASNI. The definitions are grounded in the functional necessities imposed by the phenomenon under investigation (MPAASNI) and situated within the broader philosophical landscape by acknowledging related concepts and problems, using accessible references primarily for contextual orientation rather than direct support for the functional definition itself.

2.1. Reference (R): The Capacity for Directed Content

The first component of the Triadic System is Reference (R). Functionally defined within this framework, *Reference is the capacity of a system to establish and utilize connections between its internal states or representational elements (such as concepts, symbols, or neural patterns) and specific entities, properties, relations, or states of affairs.* These targets of reference may exist internally or externally to the system, and they may be tangible (physical objects, events) or non-tangible (abstract numbers, concepts, fictional entities, possibilities). The indispensable functional role of Reference within any system capable of MPAASNI is to provide *directed content*. It is the capacity that enables the system's internal states to be *about* something determinate, to possess intentionality in the sense of being directed towards objects or situations distinct from the states themselves. Without this capacity, the system's internal states would remain isolated, lacking the semantic link to anything beyond their own intrinsic properties or syntactic relations, thus rendering impossible the formation of attitudes *about* the world, abstract domains, or even other internal states conceived as distinct objects of thought.

To elaborate on this functional definition:

- *"Establish and utilize connections"*: This points to the operational aspect of Reference. It is not merely a static property but involves processes by which links are formed (e.g., through perception linking a visual input to a concept, through definition linking a term to a concept, through stipulation linking a symbol to an object) and subsequently used in further cognitive operations (e.g., accessing information associated with the referent, combining referential elements into complex structures). The system must be able not

only to form these links but also to activate and manipulate them as part of its processing.

- "*Internal states or representational elements*": This acknowledges flexibility regarding the specific nature of the internal vehicles of Reference. The framework remains neutral on whether these are linguistic symbols in a "language of thought," distributed patterns in a connectionist network, or some other form of internal state, as long as these states can functionally bear the referential connections. The crucial aspect is the *functional* role of these states in mediating the connection to the referent.
- "*Specific entities, properties, relations, or states of affairs*": This defines the range of potential targets or *referents*. The capacity for Reference must accommodate the diversity of things that propositional attitudes can be about, including concrete particulars ("that cat"), kinds ("cats"), properties ("redness"), abstract objects ("the number 7"), relations ("being on top of"), complex events ("the fall of Rome"), possibilities ("the possibility of rain"), and potentially even the system's own internal states when reflected upon ("my belief that..."). The defining characteristic is that the referent is treated as distinct from the referring state itself.
- "*Directed content*": This encapsulates the core functional output of Reference necessary for MPAASNI. Reference imbues internal states or structures with semantic content that points beyond themselves. It provides the "building blocks" – the concepts of objects, properties, and relations – that can be assembled into structured propositions. This directedness is what allows a belief to be *about* the weather, or a desire to be *about* a future state. It provides the semantic grounding without which inference would be mere symbol manipulation and preference would lack specific objects.

The functional necessity of Reference, defined in this way, stems directly from the nature of MPAASNI. Propositional attitudes are *about* things. They require content that refers. A belief must be a belief *that P*, where P is a proposition composed of referring elements. A desire must be a desire *that Q*, where Q is a proposition composed of referring elements. Without the capacity to form these referentially grounded contents, the very notion of a propositional attitude directed towards something specific dissolves. Therefore, any system capable of MPAASNI must possess the functional capacity described here as Reference.

It is crucial to distinguish this functional definition from specific theories proposing *mechanisms* for Reference. Philosophical debates concerning how reference is achieved – whether through descriptive content associated with a term (descriptivism, influenced by Frege and Russell, see Lycan, *Philosophy of Language*, Ch. 2-3), through causal-historical chains linking term usage back to an initial baptism (causal theories, associated with Kripke, *Naming and Necessity*, 1980, and Putnam, "The Meaning of 'Meaning'," 1975), through direct acquaintance, or through other means – concern the implementation or grounding of the referential connection. Our functional definition remains neutral on these mechanisms. It specifies *what* the system must be able to achieve functionally (establish directed content) for MPAASNI to be possible, regardless of *how* it achieves it. The existence of these diverse and complex theories highlights the philosophical significance and difficulty of explaining the referential capacity, but the functional requirement for *some* such capacity providing directed content remains a necessary presupposition for MPAASNI. The grounding of this definition lies in the logical requirement that a state cannot be meaningfully *about* something without a functional link enabling that "aboutness."

2.2. Inference (I): The Capacity for Systematic Processing and Relation of Contents

The second component of the Triadic System is Inference (I). Functionally defined within this framework, *Inference is the capacity of a system to process and relate contents (typically referentially constituted propositional contents), deriving new contents from existing ones or establishing structural relationships between them according to systematic patterns or rules.* The indispensable functional role of Inference within any system capable of MPAASNI is to provide *systematic structure and processing*. It is the capacity that enables the system to move beyond isolated propositions, to connect beliefs with evidence, to evaluate consistency, to derive implications, to formulate explanations, to engage in reasoning, and to structure complex bodies of knowledge or plans. Without this capacity, the system might possess referentially grounded contents, but these contents would remain largely disconnected and unorganized, lacking the dynamic interrelations and structured manipulation required for complex thought and the functional significance of attitudes within a coherent system.

To elaborate on this functional definition:

- *"Process and relate contents"*: This highlights the transformative and relational function of Inference. It operates *on* the contents provided by Reference (and potentially other inputs), changing the system's informational state by generating new contents or by establishing explicit links (e.g., implication, support, inconsistency) between existing contents.
- *"Deriving new contents from existing ones"*: This covers the generation of conclusions from premises, characteristic of deductive, inductive, and abductive reasoning patterns. It allows the system to extend its reach beyond immediately given information.
- *"Establishing structural relationships"*: This includes identifying logical connections (like entailment or contradiction), probabilistic dependencies, causal links (as inferred), or coherence relations among propositions. This function organizes the system's contents into structured networks rather than mere collections.
- *"According to systematic patterns or rules"*: This emphasizes that Inference is not arbitrary but follows regular procedures. These patterns might correspond to logical rules (like Modus Ponens), probabilistic principles (like Bayesian updating), heuristics (like inference to the best explanation based on criteria like simplicity), or other systematic ways of transforming information. The system must be able to apply these patterns reliably.

The functional necessity of Inference, defined in this way, stems from the requirements of MPAASNI. Propositional attitudes are not isolated states; they exist within complex networks. Beliefs support or contradict other beliefs; beliefs and desires jointly lead to intentions; hopes are based on inferences about possibilities. The *meaningfulness* (functional significance) of an attitude often depends on its inferential connections – a belief is significant because it can serve as a premise for further reasoning or guide action based on inferred consequences. Justification for beliefs heavily relies on inferential support from evidence or other beliefs. Evaluating the rationality of a set of attitudes inherently involves assessing their inferential coherence. Therefore, any system capable of MPAASNI must possess the functional capacity described here as Inference to structure, relate, and process the contents provided by Reference in these systematic ways.

This functional definition must be distinguished from specific formal logical systems (like classical propositional logic, predicate logic, modal logic, probabilistic logics) or particular psychological models of human reasoning. While formal logic provides precise descriptions of certain valid inference patterns (particularly deductive ones, see Sainsbury,

Logical Forms: An Introduction to Philosophical Logic, Second Edition, 2001), the functional capacity for Inference within the Triadic System is broader, encompassing the systematic patterns underlying inductive and abductive reasoning as well, which are crucial for empirical knowledge and explanation (see Audi, *Epistemology: A Contemporary Introduction to the Theory of Knowledge*, Third Edition, 2010, Ch. 7-8 on inductive reasoning and inference to the best explanation). The definition focuses on the necessary functional role of *systematic content processing and relation*, regardless of the specific algorithms or heuristics employed by the system. The grounding of this definition lies in the logical requirement that meaningful, interconnected attitudes, and the possibility of reasoning or justification concerning them, necessitate a capacity for systematic processing and relation of their contents.

2.3. Preference (P): The Capacity for Internal Stance, Weighting, and Selection

The third component of the Triadic System is Preference (P). Functionally defined within this framework, *Preference is the capacity of a system to exhibit an internal stance, weighting, orientation, or selection regarding contents, states of affairs, or processing pathways*. The indispensable functional role of Preference within any system capable of MPAASNI is twofold: first, to provide the specific *mode* of the attitude (distinguishing belief from desire, hope from fear, etc.) by constituting the system's particular orientation towards the propositional content; and second, to imbue the attitude with *functional significance* or *meaningfulness* by assigning internal weights, priorities, or value-orientations that influence how the content is processed, integrated, maintained, and utilized in guiding the system's overall state and potential interactions. Without this capacity, the system might process referentially grounded contents through inference, but it would lack the internal orientation necessary to *hold* an attitude in a specific mode (e.g., believing vs. desiring) and the internal weighting necessary for that state to play its characteristic functional role within the system's economy.

To elaborate on this functional definition:

- *"Internal stance, weighting, orientation, or selection"*: This captures the diverse ways Preference manifests functionally. It can be a disposition to treat a proposition as true (belief stance), a disposition towards realizing a state of affairs (desire orientation), a weighting assigned to evidence or hypotheses based on epistemic criteria (simplicity, coherence), a selective focus on certain types of information or inferential pathways due to cognitive tendencies or practical goals, or an affective coloring influencing evaluation (fear, hope). It represents the system's internal "tuning" or "disposition" concerning information and processing.
- *"Regarding contents, states of affairs, or processing pathways"*: Preference operates at multiple levels. It applies to the propositional contents themselves (preferring P to be true, preferring state Q to be realized), to the states of affairs referred to (valuing certain outcomes), and also to the cognitive processes involved (preferring certain inferential methods, prioritizing certain referential targets based on salience).
- *"Providing the specific mode of the attitude"*: Different attitudes involve different preferential orientations towards the same propositional content. Believing P involves a preference aligned with P's truth. Desiring P involves a preference aligned with P's realization. Fearing P involves a preference aligned with P's non-realization (or avoidance). The Preference capacity enables these distinct modes of relating to content.
- *"Imbuing the attitude with functional significance/meaningfulness"*: The weighting or orientation provided by Preference determines how an attitude influences other states

and processes. A strongly held belief (high preference weighting for its truth) will have more impact on inference and action than a weakly held one. A strong desire (high preference weighting for realization) will more strongly motivate planning and action. This internal weighting is crucial for the attitude's role in the system's economy.

The functional necessity of Preference, defined this way, arises from the nature of MPAASNI. An attitude is not just structured content; it is content held *in a certain way*. The "attitude" part of "propositional attitude" necessitates an internal stance or mode. The "meaningful" part necessitates functional significance derived from internal weighting or orientation. A system must be able to differentiate its relation to various contents (believing some, desiring others) and assign varying degrees of importance or commitment to them for these states to function as beliefs, desires, etc., in a recognizable sense. Therefore, any system capable of MPAASNI must possess the functional capacity described here as Preference to provide this internal stance and functional weighting.

This functional definition is distinct from specific theories of value (e.g., realism vs. anti-realism about moral values), theories of emotion, or detailed models of motivation or decision-making (though decision theory, focusing on rational preference and utility, provides a formal framework for a subset of preference-related phenomena, see Resnik, *Choices: An Introduction to Decision Theory*, 1987). The functional definition focuses solely on the necessary role of *internal stance, weighting, orientation, or selection* in constituting the attitude mode and providing functional significance, regardless of the specific source or nature of these preferences (whether learned, innate, rational, irrational, affective, or purely cognitive like epistemic values). The grounding of this definition lies in the conceptual requirement that an "attitude" implies a stance and that "meaningfulness" implies functional weighting within a system.

In summary, this section has provided functional definitions for Reference (R), Inference (I), and Preference (P) as the necessary constitutive capacities for the possibility of MPAASNI. Reference provides directed content, Inference provides systematic processing and relation of content, and Preference provides the internal stance/weighting that determines the attitude mode and functional significance. Each definition focuses on the functional role required by the nature of MPAASNI, grounded in the structure of the phenomenon itself and related philosophical concepts, while remaining neutral on specific implementation mechanisms. These definitions establish the conceptual components whose necessary interrelation – the Trialectics – will be analyzed in the next section as the functional architecture enabling the emergence of meaningful propositional attitudes about something not itself.

3. Articulating the Trialectics

3.1. From Components to Interrelation

The preceding section delineated the functional necessities of Reference (R), Inference (I), and Preference (P) as distinct capacities required for any system capable of exhibiting Meaningful Propositional Attitudes About Something Not Itself (MPAASNI). Reference provides directed content, enabling states to be *about* something. Inference provides systematic processing, enabling the formation, manipulation, and relation of structured contents. Preference provides the internal stance and weighting, enabling the system to adopt specific attitude modes (belief, desire, etc.) towards contents and giving these attitudes functional significance. However, merely establishing the necessity of these three capacities existing

independently within a system is insufficient to account for the integrated nature of MPAASNI. A meaningful belief is not simply a piece of referential content existing alongside unrelated inferential processes and unattached preference states. It is a single, unified state where the directed content *is* the content processed inferentially and *is* the content towards which a specific preferential stance (acceptance as true) is held, and where these aspects are functionally interconnected. Similarly, a meaningful desire involves specific referential content, inferential processing related to its realization, and a preferential stance favoring that realization, all bound together.

The philosophical challenge, therefore, shifts from identifying the necessary components to understanding their necessary mode of *interrelation*. How must these three functional capacities (R, I, P) connect and interact within a system to constitute the specific, integrated phenomenon of MPAASNI? Our descriptive analysis suggests that the required interrelation is not merely additive or sequential, where one capacity simply provides input to the next in a linear fashion. Instead, the evidence points towards a more complex dynamic of *mutual conditioning*, where the state and operation of each capacity are continuously influenced by, constrained by, and constitutive of the state and operation of the other two. It is this specific mode of complex, interdependent functioning that this framework designates as the *Trialectics*. This section aims to articulate the nature of this Trialectics by meticulously describing the specific forms of mutual conditioning that obtain among Reference, Inference, and Preference when a system exhibits MPAASNI. This description forms the core of the proposed functional architecture.

3.2. Defining the Trialectics: A Mode of Mutual Conditioning

The term *Trialectics* is employed here to denote the specific functional mode of operation within a system characterized by *systematic and constitutive mutual conditioning among its capacities for Reference, Inference, and Preference*. It is not a separate entity or capacity but describes the way R, I, and P operate together when integrated to support MPAASNI. The term combines 'tri-' (three) with 'dialectics' (evoking dynamic interaction, interdependence, and mutual influence) to capture the essence of this three-way functional relationship.

- **Mutual Conditioning:** This is the defining feature. It means that the state of R is not independent of the states of I and P; the state of I is not independent of the states of R and P; and the state of P is not independent of the states of R and I. Changes or specific configurations in one capacity have direct implications for the functioning and potential states of the others within the system. This goes beyond simple input-output relationships; it suggests a deeper level of functional entanglement.
- **Systematic:** The conditioning is not random or chaotic but occurs according to regular patterns inherent in the functional architecture. The way P guides I, or I establishes objects for R, follows discernible principles related to the functions of each capacity.
- **Constitutive:** The claim is stronger than mere influence. The Trialectics describes the functional interrelation that *constitutes* the possibility of MPAASNI. The integrated state of a meaningful belief or desire *is* the state of R, I, and P operating in this specific, mutually conditioning configuration with respect to a particular propositional content. The attitude emerges *from* this specific dynamic interplay.

Therefore, the Trialectics is the description of a functional state where R, I, and P are operating not as separate modules but as deeply intertwined aspects of a single, integrated process supporting complex cognitive states directed towards the world.

3.3. Describing the Mutual Conditioning in Detail

To articulate the Trialectics fully, we must detail the specific pathways of mutual conditioning among R, I, and P, building upon the initial sketches provided earlier. This involves systematically describing how each capacity conditions, and is conditioned by, the other two, grounding these descriptions in the functional roles defined in Section 2.

3.3.1. Conditioning Relations Involving Reference (R)

- **(I → R) Inference conditions Reference:**
 - *Content Generation:* Inferential processes (particularly abduction and deduction from theories) can postulate the existence of entities or properties not directly observed (e.g., theoretical entities in science, implications of mathematical axioms). These inferred posits require the system to establish new referential links – terms or concepts that refer to these inferred entities (e.g., referring to 'electron', 'set', 'unconscious motivation'). Inference thus populates the domain of potential referents.
 - *Reference Refinement:* Inference can refine the understanding of what an existing term refers to. As a system draws more inferences about an object or kind (e.g., inferring its internal structure, causal properties, or historical origin), the referential link associated with the term for that object/kind becomes more precise or its conditions of application become clearer. The inferential network surrounding a concept contributes to fixing its reference.
 - *Disambiguation:* When referential ambiguity occurs (e.g., a term could refer to multiple things), inference based on context or background knowledge is often necessary to determine the intended referent in a specific instance. The inferential process selects or clarifies the operative referential link.
- **(P → R) Preference conditions Reference:**
 - *Attentional Focus & Salience:* Preference structures (driven by practical goals, epistemic interests, or cognitive biases) guide the system's attention, making certain aspects of the environment or conceptual space more salient. This selective attention influences which entities or properties become targets for the establishment of referential links. We develop vocabulary and concepts (referential tools) for things that matter to us (reflecting preference).
 - *Categorization & Framing:* Preferences can influence how the system categorizes objects or frames situations, thereby affecting the conceptual structures used for reference. A preference for seeing similarities might lead to broader categories and referential terms, while a preference for distinctions might lead to finer-grained ones. The way something is referred to (e.g., "collateral damage" vs. "civilian deaths") can be guided by underlying preferences or rhetorical goals.
 - *Referential Standard Setting:* Epistemic preferences (e.g., for precision, clarity) can influence the standards the system adopts for successful reference. A high preference for precision might lead to stricter requirements for identifying referents or disambiguating terms.

- **(R → I) Reference conditions Inference:**
 - *Provides Operands:* Referentially constituted propositions are the necessary inputs or operands upon which inference patterns operate. Inference requires structured content that is *about* something. The range and nature of what the system can refer to thus determines the domain over which its inferences can range.
 - *Determines Applicability of Rules:* The specific referents or types of entities involved in a proposition can determine which inference rules are applicable. Inferences involving quantification depend on the nature of the domain referred to. Inferences involving specific relations depend on the properties of the referents that stand in those relations.
 - *Grounds Truth Evaluation:* The success of reference is crucial for evaluating the truth of premises used in inference, particularly for inferences intended to yield knowledge about the world. An inference, however formally valid, cannot yield truth about the world if its premises fail to refer accurately.
- **(R → P) Reference conditions Preference:**
 - *Provides Objects of Preference:* Entities, properties, or states of affairs must be capable of being referred to (identified, conceptualized) before they can become objects of preference. We desire, fear, or value things we can think or talk about. The system's referential capacity limits the set of potential objects for its preference structures.
 - *Informs Preference Formation:* The properties of a referent, as understood through referentially grounded beliefs, inform the preferences formed towards it. Learning (via R and I) that an object is beneficial or harmful directly influences preferences regarding it.
 - *Influences Preference Intensity:* The way something is referred to or represented (e.g., vividly vs. abstractly) can influence the intensity of the preference felt towards it.

3.3.2. Conditioning Relations Involving Inference (I)

- **(R → I) Reference conditions Inference:** (Already described above: provides operands, determines rule applicability, grounds truth evaluation).
- **(P → I) Preference conditions Inference:**
 - *Goal Direction & Relevance:* Practical preferences (goals, desires) direct inferential processes towards pathways relevant to achieving those goals. We infer means to ends, predict outcomes relevant to our desires, and solve problems defined by our preferences.
 - *Epistemic Guidance:* Epistemic preferences (for simplicity, coherence, explanatory power, empirical adequacy) guide the selection and evaluation of hypotheses in abduction and induction, and influence the acceptance criteria for conclusions. They act as non-formal rules or criteria guiding the application of formal or informal inference patterns.
 - *Processing Strategy Selection:* Preferences (including cognitive biases) can influence the choice of inferential strategies (e.g., preferring heuristic shortcuts over computationally expensive algorithms, preferring inferences that confirm existing beliefs).
 - *Termination Criteria:* Preferences can influence when an inferential process is terminated (e.g., stopping when a sufficiently 'good' explanation is found).

according to preferred criteria, even if not logically exhaustive; stopping practical reasoning when a satisfactory plan is formed).

- **(I → R) Inference conditions Reference:** (Already described above: content generation, refinement, disambiguation).
- **(I → P) Inference conditions Preference:**
 - *Consequence Evaluation:* Inference allows the system to determine the likely consequences of potential actions or states of affairs. These inferred consequences then inform the evaluation of those actions or states according to existing preference structures, potentially leading to the revision or refinement of preferences.
 - *Preference Coherence Check:* Inference can be used to check the logical consistency or coherence among a system's preferences (e.g., detecting if preferring A over B, B over C, and C over A violates transitivity). This inferential analysis can prompt preference adjustment.
 - *Deriving Instrumental Preferences:* Inference (practical reasoning) identifies necessary means to achieve intrinsically preferred ends. This process derives instrumental preferences (preferences for the means) from intrinsic preferences (preferences for the ends).
 - *Justification of Preferences:* While the ultimate grounds of preference might be non-inferential, arguments (inferences) can be constructed to support or challenge the holding of certain preferences based on their relation to other accepted values or beliefs.

3.3.3. Conditioning Relations Involving Preference (P)

- **(R → P) Reference conditions Preference:** (Already described above: provides objects, informs formation, influences intensity).
- **(I → P) Inference conditions Preference:** (Already described above: consequence evaluation, coherence check, deriving instrumental preferences, justification).
- **(P → R) Preference conditions Reference:** (Already described above: attentional focus, categorization, standard setting).
- **(P → I) Preference conditions Inference:** (Already described above: goal direction, epistemic guidance, strategy selection, termination criteria).

This detailed mapping of the mutual conditioning pathways illuminates the intricate nature of the Trialectics. It shows that Reference is not merely about labeling but is guided by priorities (P) and refined by reasoning (I). Inference is not just mechanical derivation but operates on meaningful content (R) and is directed by goals and criteria (P). Preference is not arbitrary inclination but is typically about identifiable things (R) and informed or constrained by reasoning (I). The Trialectics is the state where these pathways of influence are active and interdependent, forming a functionally integrated system.

3.4. The Trialectics as Functional Architecture

Describing the Trialectics in this way characterizes it as a *functional architecture*. It specifies the necessary functional components (R, I, P) and, crucially, the necessary pattern of dynamic interrelations (mutual conditioning) among them. It is an abstract description of *how* a system must be organized *functionally* to support MPAASNI. This architecture is defined by the flow of influence and dependency among its parts, rather than by specific physical

structures or algorithms (which would constitute an implementation of this architecture). Different physical systems (e.g., biological brains, potentially sophisticated AI, structured collectives) might implement this functional architecture in vastly different ways, but according to the thesis, any system capable of MPAASNI must instantiate this pattern of triadic mutual conditioning at the functional level. The Trialectics, therefore, represents the necessary abstract organizational blueprint for meaningful, directed, propositional thought and attitude.

This detailed articulation of the Trialectics provides the crucial conceptual link between the basic functional capacities (R, I, P) and the complex cognitive phenomenon they enable (MPAASNI). The next step is to formally define MPAASNI based on its functional requirements, setting the stage for the core argument that demonstrates why the Trialectic architecture, as described here, is its necessary condition of possibility.

4. Defining the Phenomenon: Meaningful Propositional Attitudes About Something Not Itself (MPAASNI)

Having meticulously defined the three fundamental functional capacities of Reference (R), Inference (I), and Preference (P), and having articulated the Trialectics as the specific mode of their functional interrelation characterized by systematic and constitutive mutual conditioning, the analysis now focuses on providing an equally precise and functionally grounded definition of the target phenomenon for which the Trialectics is claimed to be necessary: *Meaningful Propositional Attitudes About Something Not Itself (MPAASNI)*. This definition is crucial, as the argument for necessity hinges upon showing that the inherent functional requirements of MPAASNI are precisely those met by the Trialectic architecture. The definition will be constructed incrementally by analyzing each component of the phrase, grounded in the functional roles identified previously, and aiming for maximal descriptive clarity without introducing normative or implementation-specific details.

4.1. Deconstructing the Term: Components of MPAASNI

The phrase "Meaningful Propositional Attitudes About Something Not Itself" encapsulates four key characteristics that distinguish this class of cognitive states:

- They are Attitudes.
- They are Propositional.
- They are About Something Not Itself (directedness/intentionality).
- They are Meaningful (functionally significant).

A precise functional definition requires elaborating the necessary contribution of each characteristic to the overall phenomenon.

4.2. The "Attitude" Component: Internal Stance or Mode

The term "Attitude" signifies that these are not passive representations or mere computational outputs, but involve a specific internal stance, orientation, or mode of relation adopted by the system towards a given content. As explored in the definition of Preference (P), this internal stance is what distinguishes believing *that* P from desiring *that* P, hoping *that* P, fearing *that* P, intending *that* P, etc. Each distinct attitude type corresponds to a different functional mode of relating to the propositional content. Believing involves a stance of acceptance as true or representing actuality. Desiring involves a stance oriented towards

realization or bringing about. Fearing involves a stance oriented towards avoidance or negative evaluation of potential realization. Intending involves a stance of commitment towards bringing about. Functionally, this "attitude" component necessitates a capacity within the system to assign different internal weightings, priorities, or orientations to contents, which determines how those contents function within the system's overall economy. This corresponds directly to the functional role identified for Preference (P). Therefore, a necessary condition for a state to be an "Attitude" in this sense is the involvement of the Preference capacity to provide this internal mode or stance.

4.3. The "Propositional" Component: Structured Content

The term "Propositional" specifies the *type* of content towards which the attitude is directed. Propositions, within this framework, are understood functionally as structured, truth-evaluable contents. They are composed of constituent elements (concepts, representations of objects, properties, relations) combined according to certain structural principles (often analogous to logical or grammatical structure) such that the entire complex content can, in principle, be assessed for truth or correspondence to a state of affairs. Believing "The cat is on the mat" involves relating the concepts 'cat', 'mat', and 'on' in a specific structure that corresponds to a potential state of affairs which is either true or false. This structured content allows propositions to serve as premises and conclusions in reasoning, to be negated, quantified over, and embedded within more complex propositions. Functionally, the capacity to form and entertain such structured, truth-evaluable contents necessitates a mechanism for combining basic referential elements into complex representations that mirror potential states of affairs. This requires both the capacity to establish the basic referential elements (Reference (R)) and the capacity to structure these elements according to systematic combinatorial principles, often involving relational concepts also established via R, and manipulated or analyzed via Inference (I) (e.g., understanding the logical structure of the proposition). The "Propositional" component thus necessitates the involvement of both R (for the constituent elements) and I (for the structure and potential manipulation).

4.4. The "About Something Not Itself" Component: Directedness/Intentionality

This crucial qualifier emphasizes the directedness or intentionality of the attitude. The propositional content is not merely an internal configuration but purports to be *about* or *directed towards* entities, properties, relations, or states of affairs that are distinct from the attitude state itself, and often distinct from the system's internal processing altogether. The 'something not itself' can be part of tangible reality (a physical object, an event), an abstract domain (mathematical objects, concepts), a potential future state, a fictional world, or even another mental state (in higher-order attitudes). Functionally, this directedness necessitates a capacity within the system to establish links between its internal propositional structures and these distinct target domains. This capacity is precisely the functional role identified for Reference (R). Reference provides the semantic connection, the "aboutness," linking the internal representation (the proposition) to its purported subject matter. Without the Reference capacity operating to establish this directed link, the propositional structure would remain purely internal syntax, failing the "about something not itself" criterion.

4.5. The "Meaningful" Component: Functional Significance

The term "Meaningful" denotes that these propositional attitudes are not inert representations but possess functional significance within the system's overall cognitive and

potentially practical economy. They play roles in guiding subsequent thought, deliberation, planning, action, and interaction with the environment. A belief influences behavior based on perceived facts; a desire motivates goal-directed activity; an intention organizes future actions. This functional significance arises from the *integration* of the attitude within the system's broader network of states and processes. It depends on the attitude's content (R) being available for processing (I), the attitude's specific mode (P) determining its influence, and the connections established through reasoning (I) linking it to other attitudes (beliefs, desires) and potential actions. The meaningfulness is thus a product of the attitude's embeddedness and causal efficacy within the system, which requires the coordinated operation of R (providing the content to be significant), I (processing the content and linking it to consequences/actions), and P (providing the internal stance/weighting that determines *how* it influences the system). Therefore, the "Meaningful" component necessitates the integrated functioning of R, I, and P – specifically, the dynamic interplay described by the Trialectics.

4.6. Synthesizing the Definition: Necessary Functional Requirements for MPAASNI

Combining these analyses, we arrive at a functional definition of Meaningful Propositional Attitudes About Something Not Itself (MPAASNI) based on their necessary constitutive requirements:

An MPAASNI is a state of a system characterized by:

1. *Directed Propositional Content*: The state involves a structured, truth-evaluable content (a proposition) whose constituent elements are functionally linked (via Reference) to specific entities, properties, relations, or states of affairs distinct from the attitude state itself.
2. *Systematic Processability*: This propositional content is embedded within the system such that it can be systematically processed, related to other contents, and used in reasoning patterns (via Inference).
3. *Specific Internal Stance and Functional Weighting*: The system adopts a specific internal mode of orientation or weighting towards this propositional content (via Preference), determining the attitude type (belief, desire, etc.) and endowing the state with functional significance within the system's overall cognitive/practical economy.
4. *Functional Integration*: These three aspects – directed content, systematic processability, and internal stance/weighting – are not merely co-present but are functionally integrated, meaning the content provided by R is the content processed by I and oriented towards by P, and the operations of I and P are conditioned by the specific content from R and the specific stance from P, respectively. This integration gives the state its unified character and functional role.

This definition highlights that MPAASNI are complex, integrated states requiring the simultaneous and interconnected operation of capacities for Reference, Inference, and Preference. The necessity of each capacity arises directly from the defining features of the phenomenon itself (directed content requires R, systematic processing requires I, internal stance requires P). The requirement for functional integration (d) points directly towards the necessity of the Trialectics, as the Trialectics is precisely the description of this integrated, mutually conditioning mode of operation among R, I, and P.

This detailed definition of MPAASNI, grounded purely in the functional requirements necessary to constitute such states as described, provides the specific target phenomenon. The

next section will present the core argument demonstrating why the Trialectic architecture, as articulated in Section 3, is the necessary condition for the possibility of states meeting this definition.

5. The Argument for Necessity: Trialectics as the Ground for MPAASNI

5.1. Linking Architecture to Phenomenon: The Core Claim

Having defined the functional capacities of Reference (R), Inference (I), and Preference (P) in Section 2, articulated the Trialectics as the specific mode of their functional interrelation characterized by mutual conditioning in Section 3, and provided a functional definition of Meaningful Propositional Attitudes About Something Not Itself (MPAASNI) based on their necessary constitutive requirements in Section 4, we now proceed to the central argument of this paper. The argument aims to demonstrate, with maximal conceptual rigor and leaving no identified gaps within the framework, the validity of the thesis: *The Trialectic Functioning of a System is a Necessary Condition for the Possibility of that System Having Meaningful Propositional Attitudes About Something Not Itself*.

The core claim connects the defined functional architecture (the Trialectics) to the defined cognitive phenomenon (MPAASNI). It asserts that the specific, integrated functional structure described by the Trialectics is not merely one way among others that might support MPAASNI, but is, within the conceptual space defined by our functional analysis, the *only* way the necessary requirements for MPAASNI can be simultaneously met and integrated. The argument proceeds by demonstrating that the absence of the Trialectic interrelation logically precludes the possibility of the functional integration identified in Section 4.6(d) as essential for constituting MPAASNI.

5.2. The Necessity Argument: Demonstrating Dependence via Conceptual Analysis

The argument relies on establishing a necessary connection between the definition of MPAASNI and the definition of the Trialectics. It proceeds through the following steps, based entirely on the previously established functional definitions:

Premise 1 (from Section 4.6): A state qualifies as an MPAASNI if and only if it possesses the integrated features of:

1. Directed Propositional Content (functionally requiring Reference).
2. Systematic Processability (functionally requiring Inference).
3. Specific Internal Stance and Functional Weighting (functionally requiring Preference).
4. Functional Integration of (a), (b), and (c) (meaning the content from R is the content processed by I and oriented towards by P, and the operations of I and P are conditioned by the specific states of R and P/R respectively).

Premise 2 (from Section 3.2 & 3.3): The Trialectics is defined as the functional mode of operation within a system characterized by systematic and constitutive mutual conditioning among its capacities for Reference, Inference, and Preference, such that:

- R conditions and is conditioned by I and P.
- I conditions and is conditioned by R and P.
- P conditions and is conditioned by R and I.

(The specific pathways of conditioning are detailed in Section 3.3).

Argument Step 1: Showing that the Constitutive Requirements of MPAASNI (a, b, c) Necessitate the Presence of R, I, P Capacities.

This was established in Section 4. The very definition of MPAASNI involves directed content (requiring R), systematic processability (requiring I), and internal stance/weighting (requiring P). Therefore, any system exhibiting MPAASNI must necessarily possess the functional capacities for Reference, Inference, and Preference. This step is foundational but insufficient on its own.

Argument Step 2: Showing that the Functional Integration Requirement (d) of MPAASNI Necessitates the Mutual Conditioning Described by the Trialectics. This is the crucial step. Requirement (d) states that for a state to be an MPAASNI, the R, I, and P components contributing to it must be functionally integrated. Let's analyze what this integration functionally entails and why it necessitates the Trialectics:

- *Integration requires that the content provided by R is available to and processed by I.* Functionally, this means the referential links establishing propositional content must serve as the operands for the inferential patterns. This implies a conditioning relation: $R \rightarrow I$ (Reference conditions Inference by providing its input).
- *Integration requires that the content provided by R is the object towards which the stance/weighting of P is directed.* Functionally, this means the preference structures must operate *on* or *with respect to* the specific, referentially determined content. This implies a conditioning relation: $R \rightarrow P$ (Reference conditions Preference by providing its object).
- *Integration requires that the processing by I operates systematically on the relevant content.* For this processing to be systematic within a cognitive economy (as required by "meaningful"), it cannot operate in a vacuum. It must be directed towards relevant content and potentially guided by criteria or goals. Functionally, this means Inference must be conditioned by the content provided ($R \rightarrow I$) and potentially guided by the system's internal orientation ($P \rightarrow I$). The necessity of P conditioning I for *meaningful* integration arises because unguided, undirected inference would not connect systematically to the system's overall functional state or goals, which are represented by P. Inference needs guidance (from P) regarding relevance, selection of pathways, or evaluation criteria to be functionally integrated in a meaningful way.
- *Integration requires that the internal stance/weighting provided by P is assigned specifically to the content provided by R, often as evaluated or modified by I.* Functionally, this means Preference must be conditioned by the specific content ($R \rightarrow P$) and potentially informed by the results of processing that content ($I \rightarrow P$). The stance needs to be *about* the right content, and its strength or nature might depend on inferential evaluation (e.g., strength of belief based on inferred evidence).
- *Integration requires that the entire system functions coherently.* This implies feedback loops. For instance, the outcomes of inferential processes (I) guided by preferences (P) might lead to the formation of new concepts or refinement of existing ones, thus altering the referential landscape ($I \rightarrow R$). Similarly, the system's preferences (P) might themselves be refined based on inferential evaluation ($I \rightarrow P$) or focused based on

referential possibilities ($R \rightarrow P$). Preferences might also guide the focus of reference ($P \rightarrow R$) or the selection of inference rules ($P \rightarrow I$).

If we assemble these necessary conditioning pathways required for the *functional integration* specified in Premise 1(d), we find precisely the set of mutual conditioning relations defined as the *Trialectics* in Premise 2:

- $R \rightarrow I, R \rightarrow P$
- $I \rightarrow R, I \rightarrow P$
- $P \rightarrow R, P \rightarrow I$

Therefore, the requirement for functional integration (d), which is necessary for a state to be an MPAASNI, logically entails the presence of the mutual conditioning among R, I, and P that defines the Trialectics.

Argument Step 3: Analyzing the Contrapositive – Absence of Trialectics Implies Absence of MPAASNI.

To further solidify the necessity claim, consider the conceptual consequence of a system possessing capacities for R, I, and P, but *lacking* the Trialectic interrelation (i.e., lacking the specified mutual conditioning).

- If R is not conditioned by I and P: Referential content is formed without being systematically refined by reasoning or guided by relevance/salience.
- If I is not conditioned by R and P: Inference operates without grounded content or guiding criteria/goals.
- If P is not conditioned by R and I: Preference states are unassigned to specific contents or unformed by reasoned evaluation.

In such a functionally disintegrated system:

- Could a state have directed propositional content? Yes, potentially, via R.
- Could this content be systematically processable? Yes, potentially, via I (though operating ungrounded).
- Could the system have internal stances/weightings? Yes, potentially, via P (though unassigned/uninformed).

However, could such a system form a *single, integrated state* that simultaneously *is* directed content *and* is systematically processed *and* embodies a specific internal stance towards *that specific content* in a *functionally significant* way? No. Because the functional links *between* R, I, and P that would bind these aspects together are, by definition, absent. The content (R) wouldn't be the specific content processed by I in a way guided by P, nor the specific content towards which P holds an informed stance. The processing (I) wouldn't be operating on functionally significant, directed content. The stance (P) wouldn't be directed at processed, referential content. The state would lack the necessary functional integration (Premise 1(d)).

Therefore, a system lacking the Trialectics, even if possessing R, I, and P capacities in isolation, cannot form states that meet the functional definition of MPAASNI. The absence of the Trialectics implies the impossibility of MPAASNI.

Conclusion of the Argument:

The analysis demonstrates that:

1. MPAASNI necessarily requires capacities for R, I, and P (Argument Step 1).
2. MPAASNI necessarily requires the functional integration of these capacities (Premise 1(d)).
3. This functional integration logically entails the presence of the specific mutual conditioning relations defined as the Trialectics (Argument Step 2).
4. The absence of the Trialectics logically precludes the functional integration required for MPAASNI (Argument Step 3).

Therefore, the Trialectic functioning of a system (as defined) is a *necessary condition* for the possibility of that system having Meaningful Propositional Attitudes About Something Not Itself (as defined). The necessity is established through strict conceptual analysis of the functional requirements inherent in the definition of MPAASNI and showing that these requirements are met only when the R, I, and P capacities operate in the mutually conditioning mode described as the Trialectics. This argument, grounded entirely within the descriptive framework established, leaves no identified conceptual gap between the definition of the phenomenon and the necessity of the proposed underlying functional architecture.

6. Scope, Particularities, and Implications

Having established the central thesis through conceptual argument—demonstrating that the Trialectic functioning (mutual conditioning of Reference, Inference, and Preference) is a necessary condition for the possibility of Meaningful Propositional Attitudes About Something Not Itself (MPAASNI)—this section explores the scope of this claim, how specific types of attitudes relate to this general necessity, and some derivative implications for understanding other complex cognitive phenomena. This exploration remains within the descriptive and analytical mode, drawing out consequences and clarifications based on the established framework.

6.1. Scope of the Thesis

The necessity derived connects a specific functional architecture (Trialectics) to a specific class of cognitive phenomena (MPAASNI). Understanding the scope involves clarifying the range of systems and situations to which this necessity applies.

- **Applicability to Systems:** The thesis applies to *any* system, regardless of its physical constitution (biological, artificial, potentially collective), that is capable of exhibiting states meeting the functional definition of MPAASNI. The necessity is functional, not tied to a particular substrate. If a system can genuinely hold beliefs or desires about things distinct from itself in a way that is functionally significant, then it necessarily instantiates the Trialectic architecture at the functional level. This includes individual cognitive agents (like humans, presumably) but could conceptually extend to sufficiently sophisticated artificial intelligences or even structured collective entities

(like the scientific community conceptually modeled earlier) if they meet the functional criteria for possessing collective reference, inference, preference, and their mutual conditioning leading to collective MPAASNI.

- **Applicability Across Object Domains:** The thesis holds irrespective of whether the "something not itself" that the attitude is about is part of tangible reality or belongs to non-tangible domains. As argued previously, attitudes towards abstract objects (numbers, concepts), possibilities, fictional entities, logical relations, or other mental states all fit the definition of MPAASNI. They require directed content (Reference, even if the mechanism differs for abstract objects), systematic processing (Inference, e.g., mathematical or logical reasoning), and an internal stance (Preference, e.g., accepting a theorem as true). Therefore, the Trialectics is necessary for the possibility of belief about mathematical truths, desire for abstract justice, hope regarding future possibilities, or fear concerning logical contradictions, just as it is for attitudes about physical objects and events. The necessity lies in the functional structure required to integrate directed content, processing, and stance, regardless of the domain the content refers to.
- **Connection to Tangible Reality (Revisited):** While applicable to non-tangible domains, the framework's connection to tangible reality remains crucial. Since many, perhaps most, of the MPAASNI exhibited by embodied agents like humans *are* directed towards tangible reality, the Trialectics is posited as the necessary functional architecture enabling our sophisticated cognitive engagement *with the physical world* through states like belief, desire, and intention concerning objects and events within it. It describes the functional bridge enabling internal states to be meaningfully directed towards the tangible.

6.2. Particular Attitudes as Configurations of the Trialectics

The thesis establishes the Trialectics as a *necessary* architecture for the *general* class of MPAASNI. It does not imply that all MPAASNI are identical. The differences between specific attitude types – believing vs. desiring vs. hoping vs. fearing, etc. – are accounted for within the framework as different configurations or specific patterns of activation and weighting within the necessary Trialectic structure. The fundamental interrelation of R, I, and P must be present for any MPAASNI, but the specific way they interact and, crucially, the nature and orientation of the Preference (P) component, determines the particular attitude type.

- **Belief Configuration:** Characterized by a Preference orientation towards truth/accuracy. This P-configuration guides R towards accurate representation and I towards evidence assessment and consistency checking, resulting in a state (belief) that functions as a map of how the world is taken to be.
- **Desire Configuration:** Characterized by a Preference orientation towards realization/bringing about. This P-configuration guides R towards representing the desired state and I towards means-end reasoning and planning, resulting in a state (desire) that functions as a motivator for action towards a goal.
- **Hope Configuration:** Characterized by a complex Preference orientation involving desire for realization coupled with specific processing of likelihood information (via I) regarding an uncertain future state (represented via R).
- **Fear Configuration:** Characterized by a Preference orientation towards non-realization/avoidance regarding a negatively evaluated potential state (represented via R), guiding I towards risk assessment and avoidance planning.

Other attitudes (intention, doubt, regret, etc.) can similarly be analyzed as involving distinct configurations of the Preference component interacting in specific ways with Reference and Inference. The Trialectics provides the necessary functional "grammar" or "operating system," while the specific settings and dynamics within that system, largely determined by the state of the Preference component, constitute the particular attitudes. The generality of the thesis lies in the necessity of the underlying grammar; the particularity lies in the specific "utterances" or states generated by different settings within that grammar.

6.3. Phenomena Impossible Without Trialectics (Derivative Impossibilities)

If the Trialectics is necessary for MPAASNI, then any phenomenon that constitutively *requires* MPAASNI as a component is also, derivatively, impossible without the Trialectics. This extends the implications of the thesis beyond individual attitudes to broader cognitive capacities and activities. Based on the analysis in Section 5, phenomena rendered fundamentally impossible (within this conceptual framework) in the absence of the Trialectic functioning include:

- **Rational Agency:** Requires beliefs and desires (MPAASNI) and practical inference operating on them.
- **Complex Deliberation and Planning:** Requires forming and manipulating beliefs and desires about potential futures (MPAASNI).
- **Meaningful Linguistic Communication (expressing/interpreting propositional attitudes):** Requires the capacity to form, express, and understand MPAASNI.
- **Systematic Knowledge Acquisition and Justification:** Requires forming justified beliefs (MPAASNI) through inference guided by epistemic preferences.
- **Higher-Order Thought (attitudes about attitudes):** Requires forming MPAASNI whose content refers to other MPAASNI.

The impossibility here is conceptual and functional: if the necessary functional architecture (Trialectics) for the component parts (MPAASNI) is absent, then the complex phenomena built upon those parts cannot arise. This highlights the foundational role attributed to the Trialectics within the framework – it is not just necessary for holding individual beliefs or desires, but for the possibility of the integrated cognitive activities that define sophisticated agency, reasoning, communication, and self-awareness, insofar as these depend on meaningful propositional attitudes directed towards a world beyond the system itself.

This exploration of scope, particularities, and implications clarifies the reach and significance of the central thesis. It applies broadly to any system capable of MPAASNI across different domains, accounts for the variety of attitude types as specific configurations within the necessary architecture, and identifies a range of complex cognitive phenomena that depend fundamentally upon this architecture.

7. Discussion and Potential Objections

7.1. Examination of Objection 1: The Necessity of Mutual Conditioning (Trialectics) vs. Sufficiency of Component Interaction

The central thesis derived from our descriptive analysis posits the Trialectic functioning—the specific mode of mutual conditioning among Reference (R), Inference (I),

and Preference (P)—as a necessary condition for the possibility of Meaningful Propositional Attitudes About Something Not Itself (MPAASNI). While the necessity of the *presence* of capacities analogous to R, I, and P for constituting such complex cognitive states might find broader acceptance, the core of the thesis lies in the claim that their specific *interrelation* in a mutually conditioning dynamic is indispensable. A significant line of potential criticism, therefore, targets precisely this aspect: Is the strong requirement of mutual conditioning, as defined by the Trialectics, genuinely necessary? Or could simpler forms of interaction among R, I, and P suffice to ground the possibility of MPAASNI? This section undertakes a dialectical examination of this objection, adhering strictly to the established descriptive framework and constraints.

The Objection – Sufficiency of Components without Full Mutual Conditioning

The objection begins by conceding, for the sake of argument, the functional definitions of R, I, and P as capacities respectively providing directed content, systematic processing, and internal stance/weighting. It also concedes that MPAASNI, as defined, involves elements corresponding to these three functions. However, the objection challenges the assertion that these functions must be integrated through the specific, complex mechanism of *mutual conditioning* described as the Trialectics.

The critic might argue as follows: "Your analysis identifies R, I, and P as necessary functional capacities. Let us grant this. You also argue that MPAASNI requires an 'integration' of these functions. Let us grant this too. However, the move from requiring integration to requiring the specific, pervasive *mutual conditioning* of the Trialectics appears overly strong and potentially unnecessary. Why must Reference be conditioned by Preference and Inference in the specific ways described for directed content to arise? Why must Inference be guided by Preference in the detailed manner outlined, beyond perhaps simple goal-setting? Why must Preference be informed by Inference and Reference so tightly?

Perhaps a simpler functional architecture could suffice. Consider a model where capacities for R, I, and P operate as largely distinct functional modules. The Reference module establishes links to the world and generates basic propositional contents. The Inference module takes these contents (and perhaps outputs from Preference) and performs derivations according to its internal rules. The Preference module assigns values or stances to certain contents or states. Integration could occur at a later stage or through simpler interfaces. For instance, the Inference module might simply take outputs from the Reference module as input, without R being significantly shaped *by* I or P. The Preference module might evaluate the outputs of I or the states represented by R, without its own structure being dynamically conditioned *by* R and I beyond receiving inputs.

Could such a system, with distinct R, I, and P modules interacting through defined input-output channels but without the deep mutual conditioning (where the internal operation and state of each is constantly shaped by the others), still give rise to states that meet the functional definition of MPAASNI? For example, couldn't a belief simply be a representation generated by R, stored in a 'belief box', evaluated as 'true' by P based on input from I processing evidence from R, without requiring that R's capacity to refer to 'evidence' was itself guided by P, or that P's criterion for 'true' was conditioned by I's processing limitations, or that I's ability to process 'evidence' was dependent on P's evaluation of sources?

The burden of proof seems to lie on demonstrating not just that R, I, and P are needed, and not just that some interaction is needed, but that the specific, complex, pervasive *mutual conditioning* described as the Trialectics is the *only* functional architecture capable of grounding MPAASNI. The objection suggests that simpler interaction models have not been sufficiently ruled out by the conceptual analysis provided, and that the Trialectics might represent an overly intricate requirement, positing more interdependency than is strictly necessary for the phenomenon it seeks to explain."

Response – The Requirement of Functional Integration for MPAASNI

The response to this objection must return to the functional definition of MPAASNI developed in Section 4, particularly the requirement for functional integration (4.6d), and demonstrate why simpler interaction models, lacking mutual conditioning, fail to meet this requirement.

The framework responds: "The objection correctly identifies the core of the thesis – the necessity of mutual conditioning, not just component presence or simple interaction. However, it underestimates the functional demands inherent in the definition of MPAASNI, particularly the 'meaningful' aspect, which necessitates deep functional integration. Let us re-examine why mere co-presence or linear input-output relations between hypothetical R, I, and P 'modules' are insufficient.

An MPAASNI, such as believing *that P*, is not just having a representation P (from R), having processed related information (via I), and having assigned a 'true' label (via P). It is a unified state where the *specific way* P is represented is suitable for the inferential processing it undergoes, where that processing *is relevantly connected* to the reasons for holding the 'true' stance, and where that stance *actively orients* the system regarding P and its implications. This requires a tighter coupling than simple modular interaction allows.

Consider the objection's proposed simpler model: R generates content, I processes it, P evaluates.

- If R operates independently, generating content without being guided by P (e.g., relevance, salience) or refined by I (e.g., theoretical context), it might produce representations that are ill-suited for meaningful inference or lack relevance to the system's concerns. The content might be technically 'directed' but functionally inert or disconnected.
- If I operates independently, receiving content from R but without being guided by P (e.g., epistemic values, goals), its processing might be purely formal or computationally exhaustive but lack direction. It might derive endless irrelevant consequences or fail to prioritize inferences needed for justification or action. Its processing would not be integrated into the system's overall functional economy.
- If P operates independently, merely evaluating outputs from R or I without its own structure being conditioned by R (the specific nature of the object of evaluation) or I (the inferential basis for evaluation), its evaluations ('true', 'desirable') would be superficial labels. Assigning 'true' without the P-structure itself being sensitive to the inferential support (I) and the referential grounding (R) specific to *that* proposition doesn't constitute the deep commitment and functional role characteristic of belief. It would be like a rubber stamp applied without regard for the document's content or origin.

The 'meaningfulness' of an MPAASNI hinges on this integration. A belief is meaningful *because* its specific referential content is precisely what is evaluated as true based on inferential support, and this 'true' stance (P) then guides further inference (I) and interaction concerning the referent (R). A desire is meaningful *because* its referential content (desired state) is linked via inference (I) to potential actions and is assigned a motivational weighting (P) that drives those inferences and actions. This requires that R, I, and P functions are not just linked by simple data passing but are mutually adjusted and conditioning. Reference needs to provide content *suited for* relevant inference and evaluation. Inference needs to *operate on* relevant content and be *guided by* evaluative criteria. Preference needs to *apply to* specific content based on *how* it relates inferentially to other states.

Therefore, the functional integration requirement (4.6d) is not met by simple modular interaction. The components must be capable of shaping each other's operation dynamically – R guided by P/I, I guided by P/R, P informed by R/I. This dynamic mutual shaping *is* the Trialectics. The necessity of the Trialectics arises directly from the necessity of this deep functional integration for constituting states that meet the full definition of MPAASNI."

Objection Refined – Appeal to a Central Executive or Workspace Model

The critic might refine their objection, acknowledging the need for integration beyond simple input-output links, but proposing an alternative architecture that achieves integration without requiring *direct* mutual conditioning between the primary R, I, and P capacities themselves.

The refinement proceeds: "Your response highlights the need for integration, particularly for the 'meaningful' aspect. Let's accept that simple modularity fails. However, consider an architecture inspired by 'global workspace' theories (e.g., Bernard Baars, *A Cognitive Theory of Consciousness*, 1988) or 'central executive' functions. In such a model, specialized R, I, and P processors might still operate with some degree of autonomy, but their outputs converge in a central workspace or are managed by an executive function. This central mechanism is responsible for integrating the information: selecting relevant referential content, initiating appropriate inferences, applying preference criteria, and broadcasting the resulting integrated state (the MPAASNI) to other systems.

In this model, the necessary integration occurs within the workspace or through the executive function, *not* necessarily through direct, pervasive mutual conditioning *between* the R, I, and P processors themselves. Reference provides content *to the workspace*. Inference operates *on content within the workspace*, perhaps directed by the executive. Preference criteria are applied *by the executive or within the workspace* to evaluate contents. The R module doesn't need its internal workings conditioned by P; it just needs to output content. The I module doesn't need its rules guided by P; it just needs to execute algorithms on workspace content when directed. The P module doesn't need to be conditioned by R and I; it just needs to provide evaluation criteria to the executive/workspace.

This architecture seems capable of achieving the functional integration required for MPAASNI – binding specific content to specific processing under specific evaluative criteria – without necessitating the strong mutual conditioning defined as the Trialectics between the core R, I, and P capacities themselves. The integration happens *centrally*, mediated by the workspace/executive, rather than *distributed* through mutual conditioning. Doesn't this

alternative model show that the specific Trialectic architecture (as pervasive mutual conditioning) is not strictly necessary?"

Response – Relocation or Insufficiency of the Central Integrator

The response must now analyze this proposed alternative architecture and argue either that it implicitly relies on the Trialectics being instantiated within the central mechanism, or that it cannot fully replicate the functional integration required for MPAASNI.

The framework responds: "The proposed central executive or global workspace model offers a potential alternative locus for integration, but a closer functional analysis reveals it either relocates the necessity of the Trialectics or fails to capture the full requirements of MPAASNI.

Consider the functions attributed to the central workspace/executive: selecting relevant referential content (R-related function), initiating appropriate inferences (I-related function), applying preference criteria (P-related function), and integrating these to form the MPAASNI. For this central mechanism to perform these integrative tasks *meaningfully*, it must itself embody the interdependencies described by the Trialectics.

- How does the executive/workspace select relevant referential content? It must do so based on current goals or criteria (Preference) and perhaps based on ongoing inferential processes (Inference). This means the mechanism performing the R-related function (selection/activation of reference) within the central unit is conditioned by P and I.
- How does it initiate appropriate inferences? It must select inference patterns relevant to the current content (Reference) and goals/criteria (Preference). The mechanism performing the I-related function (initiation/direction of inference) is conditioned by R and P.
- How does it apply preference criteria? It must apply them to specific contents (Reference) based on the results of processing (Inference). The mechanism performing the P-related function (evaluation/stance adoption) is conditioned by R and I.

Therefore, the very act of central integration, if it is to successfully bind directed content, systematic processing, and internal stance into a unified MPAASNI, requires that the central mechanism *itself* operate according to the principles of mutual conditioning among R-related, I-related, and P-related functions. The Trialectics is not eliminated; its necessity is merely relocated from the peripheral processors to the central integrator. The functional requirement for mutual conditioning remains.

Alternatively, if the central executive is conceived more simply, perhaps as just routing information between less sophisticated R, I, P modules without itself instantiating deep mutual conditioning, it's unclear how it could achieve the required functional integration. How would it ensure that the 'true' stance assigned by P genuinely reflects the inferential support (I) for that specific referential content (R), rather than being a disconnected label? How would it ensure that inferences initiated are genuinely relevant to the referential content in light of preferences? A simple routing mechanism lacks the capacity for the dynamic, content-sensitive, preference-guided mutual adjustment that constitutes the integrated state of an MPAASNI. Such a model might produce outputs that superficially resemble attitudes, but they would lack the internal coherence and functional significance derived from the integrated operation.

Thus, the central executive/workspace model either implicitly presupposes the Trialectic interrelation within its own functioning or, if conceived more weakly, fails to provide a mechanism sufficient for the functional integration necessary for MPAASNI as defined. The necessity of the mutual conditioning pattern persists."

Objection Further Refined – Questioning Universality for Simpler Attitudes

The critic might partially concede the point for highly complex, deeply integrated attitudes but question whether the *full* apparatus of mutual conditioning is necessary for *all* states that might fall under the umbrella of MPAASNI, particularly simpler or more rudimentary forms of belief or desire.

The refinement continues: "Your argument suggests that achieving the rich functional integration characteristic of, say, a carefully considered scientific belief or a complex, long-term desire might indeed require something like the Trialectic interplay. But are all MPAASNI so complex? Consider a very basic perceptual belief ('That's red') or a simple, immediate desire ('Want water'). Do these seemingly simpler states truly require the *full* feedback loop where Reference is conditioned by Inference and Preference, Inference by Reference and Preference, and Preference by Reference and Inference, all operating simultaneously and constitutively?

Perhaps simpler attitudes can be formed with only partial or unidirectional conditioning. Maybe forming the belief 'That's red' primarily involves R linking to sensory input and P assigning a 'true' stance based on perceptual reliability, with minimal involvement of complex I conditioning R or P. Maybe a basic desire like thirst involves a P state (discomfort, wanting water) directly influencing I (seeking water) and R (focusing on water-related cues), but without P itself being significantly conditioned *by* R and I in that moment.

Could it be that the Trialectics describes the structure necessary for *sophisticated* or *highly rational* MPAASNI, but not for the entire class? If simpler attitudes exist that meet the broad definition (directed content, some processing, internal stance, functional role) but are constituted by less complex or less integrated interactions among R, I, and P, then the claim that the Trialectics is necessary for *all* MPAASNI would be too strong. Perhaps necessity only applies to a subset?"

Response – Defining "Meaningful" and the Threshold for Trialectics

The response must address this challenge by returning to the qualifier "meaningful" within MPAASNI and clarifying the threshold at which the functional integration described by the Trialectics becomes necessary.

The framework concludes: "This objection raises an important point about potential gradations in cognitive complexity. However, the thesis concerns *meaningful* propositional attitudes about something not itself, and the qualifier 'meaningful' carries significant functional weight. Let us clarify the threshold implied by this.

A state qualifies as 'meaningful' in this context if it possesses functional significance derived from its integrated role within the system's cognitive economy. This implies more than just a momentary registration or reaction. It requires the state to be capable of interacting appropriately with other states, guiding behavior beyond immediate stimulus-response, serving

as a basis for further reasoning, and being subject to potential evaluation or revision within the system.

Consider the proposed simpler cases:

- **'That's red':** For this to be a *meaningful belief* rather than just a sensory classification, it must have implications (e.g., inferring it's likely a ripe fruit if associated with 'apple'), be integrated with other beliefs (e.g., consistent with believing one isn't colorblind), and be held with a stance (preference for truth) that distinguishes it from merely noticing redness. This requires linking the specific reference ('that') to the concept ('red'), processing this link via inference (classification, potential implications), and adopting a stance (preference for truth based on perception). Even here, the preference (P) for perceptual input guides the R->I process, and the inferential classification (I) refines the referential link (R) to the object's properties, while the referential link provides the object for the preference (P). The mutual conditioning, perhaps simpler, is still arguably present for the state to be a *belief* rather than raw classification. If the 'preference' aspect (the stance) were absent, it wouldn't be belief. If the 'inference' aspect (classification, minimal implications) were absent, it wouldn't be functionally integrated. If the 'reference' aspect were absent, it wouldn't be *about* the 'that'.
- **'Want water':** For this to be a *meaningful desire* rather than just a physiological drive signal, it needs propositional content directed at a state ('having water'), must be capable of guiding inference (planning how to get water), and involve a preference weighting that motivates this inference and action. The preference state (thirst/desire) guides attention (R) towards water sources and relevant inferences (I). Inferences (I) about availability inform the specific desire and plans. Reference (R) identifies potential water sources as objects relevant to the preference (P). Again, the functional integration required for this state to be a *desire* (a propositional attitude guiding action) seems to necessitate the core Trialectic interplay.

It is conceivable that systems could exhibit states with *some* characteristics of attitudes but lacking the full integration. Perhaps there are proto-beliefs or proto-desires based on simpler functional architectures. However, the thesis specifically addresses states that meet the criteria for being *meaningful propositional attitudes about something not itself*. The argument holds that any state exhibiting the full set of features—directed propositional content, systematic processability, and functionally significant internal stance, all integrated—necessarily requires the Trialectic architecture. States falling short of this integration, possibly arising from simpler interactions, would also fall short of being MPAASNI *as defined*.

Therefore, the universality of the necessity claim holds for the class of phenomena precisely defined as MPAASNI. While simpler cognitive states might exist below this threshold, the emergence of states with the specific functional profile of meaningful belief, desire, hope, fear, etc., directed towards a world beyond the system, requires the integrated, mutually conditioning dynamic of the Trialectics. The necessity is tied to the functional threshold crossed when moving from simple processing or reaction to meaningful, directed attitudes."

7.2. Examination of Objection 2: The Status of Preference as a Distinct and Necessary Component

Having addressed objections concerning the necessity of the specific *interrelation* (Trialectics), another line of potential criticism might focus on the distinctness and necessity of one of the core components itself, namely Preference (P). While the roles of Reference (for content) and Inference (for processing) might seem relatively uncontroversial as necessary elements for complex cognitive states, the inclusion of Preference as a third, distinct, and equally fundamental functional capacity necessary for MPAASNI could be questioned. Critics might argue that the functions attributed to Preference could be subsumed under Inference or Reference, or that while preference exists, its role is merely modulatory or goal-setting, rather than being a constitutive element of the attitude itself integrated via mutual conditioning.

The Objection – Preference as Subsumable or Merely Modulatory

The objection proceeds: "The framework posits Reference, Inference, and Preference as three distinct, fundamental functional capacities whose Trialectic interrelation is necessary for MPAASNI. We grant the necessity of R (for content) and I (for processing). However, the status of Preference (P) as a distinct, co-equal, and functionally necessary component in this specific way seems less secure.

Could the functions assigned to Preference – providing the internal stance/mode and functional weighting/significance – be accounted for differently?

- **Subsumption under Inference:** Perhaps the 'stance' of belief is simply the outcome of inferential processes that assign a high probability or degree of certainty to a proposition based on evidence. The 'preference for truth' might be an implicit goal embedded within the rules or algorithms of the Inference mechanism itself, aimed at maximizing predictive accuracy or minimizing cognitive dissonance. Similarly, desire might be analyzed as a type of inferential state representing a goal or a discrepancy between a current state and a target state, triggering further practical inferences. On this view, P is not a distinct capacity but a feature or output of sophisticated Inference.
- **Subsumption under Reference/Content:** Alternatively, perhaps the 'attitude mode' is encoded within the representation itself, part of the referential content. A belief representation might simply have a different format or 'tag' than a desire representation. The functional significance might derive directly from how these different types of representations are structured to interact with other parts of the system, without needing a separate 'Preference' capacity to assign weighting or stance.
- **Preference as External Modulator:** Even if preference exists as a distinct system (e.g., a motivational or value system), perhaps its role is merely to provide external goals or biases *to* the core R-I system, rather than being *mutually conditioned* by R and I and *constitutive* of the attitude state itself in the deep way the Trialectics suggests. It might set parameters for inference (e.g., 'prioritize finding food') or focus attention for reference, but the resulting belief or plan is still fundamentally constituted by R and I, with P acting as an external influence rather than an integrated component whose state is itself shaped by R and I in forming the attitude.

If any of these alternatives are viable, then the Triadic system (R, I, P) with its strong Trialectic interdependence might not be necessary. A dyadic system (R, I) perhaps modulated by external goals, or a system where P functions are realized within R or I, might suffice. The objection questions the indispensability of P as a distinct, mutually conditioned element required for *all* MPAASNI."

Response – The Irreducible Functional Role of Preference

The response must defend the distinctness and constitutive necessity of the Preference function by showing that the roles assigned to it cannot be fully captured by Reference or Inference alone, nor by treating it as a merely external modulator, without losing crucial aspects of the definition of MPAASNI.

The framework responds: "This objection challenges the decomposition of functions, specifically the status of Preference. However, a careful functional analysis reveals that the roles attributed to P – providing the specific internal stance/mode and the functional weighting/significance – are distinct from, and cannot be fully reduced to, the functions of R (directed content) or I (systematic processing), nor adequately captured by external modulation alone, if we are to account for the phenomenon of MPAASNI as described.

- **Against Subsumption under Inference:** While Inference can assign probabilities or certainty levels, this quantitative assessment is distinct from the qualitative *stance* of belief or desire. A system can infer that proposition P has probability 0.9 without necessarily *believing* P (it might suspend judgment, awaiting more evidence) or *desiring* P. Belief involves a commitment or acceptance *as true* that goes beyond mere probability calculation, reflecting an internal orientation (Preference). Similarly, desire involves an orientation towards realization that is not simply an inference about a goal state but an internal weighting favoring that state. Reducing P to I conflates calculation with commitment or motivation. Furthermore, epistemic preferences (simplicity, coherence) act as *criteria guiding* inference, suggesting they are distinct from the inferential process itself. If they were part of Inference, it becomes difficult to explain how inference rules themselves could be evaluated or selected based on these criteria.
- **Against Subsumption under Reference/Content:** Encoding the attitude mode directly into the representational format (R) faces difficulties. How would such 'tags' acquire their specific functional roles (belief-tag causing assertion-like outputs, desire-tag causing goal-pursuit outputs)? Attributing these functional roles seems to require a distinct mechanism or disposition that interprets these tags and gives them weight – precisely the function attributed to Preference. Simply having different data structures doesn't explain the difference between believing P and desiring P in terms of their internal significance and influence on the system. The 'attitude' aspect seems inherently dispositional or orientational, not purely representational.
- **Against Preference as External Modulator:** Treating Preference solely as an external input (e.g., setting goals for an R-I system) fails to capture the integrated nature of MPAASNI and the mutual conditioning described by the Trialectics. As argued in the previous objection's response, for an attitude to be *meaningful*, the internal stance (P) must be appropriately related to the specific content (R) and its processing (I). A belief's functional role depends on its truth-oriented stance being applied *to that specific content* based on its *inferential support*. A desire's role depends on its goal-orientation being applied *to that specific content* and guiding *relevant inferences*. This requires P to be sensitive and responsive to R and I. Conversely, R and I, to be functionally effective within an agent, need guidance from P regarding relevance, salience, and evaluation. If P were merely external, setting initial goals but not being dynamically conditioned by the ongoing R and I processes related to a specific attitude, the resulting state would lack the internal coherence and responsiveness characteristic of MPAASNI. For instance, how could belief revision occur effectively if the P-component (stance

towards truth) wasn't continuously sensitive to the I-component (processing of new evidence related to the specific R-content)? The mutual conditioning seems necessary for the attitude to function as an integrated, responsive state.

Therefore, the functional role of providing the internal stance/mode and the functional weighting/significance appears distinct from providing directed content (R) and systematic processing (I). This distinct functional role is indispensable for constituting the "Attitude" aspect and the "Meaningful" aspect of MPAASNI. Attempting to reduce P to R or I, or treating it as merely external, overlooks the need for an internal mechanism that assigns specific, weighted orientations to processed, directed contents in an integrated fashion. The Preference capacity fulfills this necessary role, and its integration with R and I via mutual conditioning (the Trialectics) is required for the resulting state to be a unified, meaningful propositional attitude."

Objection Refined – The Nature of Epistemic Preferences

The critic might focus specifically on *epistemic* preferences (simplicity, coherence, etc.) and argue that these, at least, seem reducible to purely inferential or logical considerations, weakening the case for P's distinctness in the context of belief formation and justification.

The refinement proceeds: "Let us grant, perhaps, that practical preferences (desires, goals) involve a distinct motivational or valuational element. But consider epistemic preferences like simplicity or coherence, which you claim guide inference ($P \rightarrow I$) and are necessary for justification within the Trialectics. Aren't these simply features of *good* inference or rational belief systems themselves? Coherence might be seen as a logical property detectable by Inference alone (checking for contradictions). Simplicity (e.g., Ockham's Razor) might be a heuristic rule *within* an inferential strategy for abduction, not an external preference guiding it. If these key 'epistemic preferences' can be re-described as properties of inference patterns or logical relations among contents (handled by I and R), then the role of a distinct P capacity in the formation of *belief* (a core MPAASNI) is diminished, potentially undermining the necessity of the full Trialectics for justified belief."

Response – Preference as Selection/Weighting Among Inferential Options

The response must distinguish between the logical properties themselves and the system's *disposition to value or select based on* those properties, arguing that the latter is the irreducible role of Preference.

The framework responds: "This refinement helpfully focuses on epistemic preferences. However, it conflates the *property* (e.g., simplicity, coherence) with the system's *functional orientation* towards that property.

- **Coherence:** While Inference (I) can detect logical inconsistency between propositions (whose content is fixed by R), the *drive* to resolve inconsistency or the *weighting* given to coherence as a reason to accept or reject beliefs is not itself part of the logical detection mechanism. Two systems could both infer an inconsistency, but one might tolerate it while the other is strongly disposed to revise beliefs to eliminate it. This disposition reflects a difference in their Preference structure regarding coherence. P

provides the *normative force* or *functional weighting* associated with the logical property detected by I.

- **Simplicity:** Similarly, while simplicity might be definable as a formal property of hypotheses or models (related to R and I), the principle that *simpler hypotheses should be preferred* (Ockham's Razor) is not itself a rule of deductive or inductive inference. It is a methodological preference, a criterion used to *select among* multiple potential conclusions generated by abductive inference (I) based on their referential content (R). Inference might generate several explanations compatible with the evidence; Preference (valuing simplicity) selects one. Without this preferential weighting, Inference alone would leave the choice undetermined or based on other criteria.

Epistemic preferences, therefore, function as selection principles or weighting mechanisms that operate *on* the outputs or processes of Inference and Reference. They represent the system's internal criteria for what counts as a 'good' explanation, a 'justified' belief, or a 'reliable' process, beyond mere logical validity or representational accuracy. These criteria are necessary for navigating situations where logic and evidence alone underdetermine conclusions (common in induction and abduction). Assigning this role of selection and weighting based on criteria like simplicity or coherence is precisely the function defined for Preference (P). Even in the epistemic domain, P remains functionally distinct from I (which executes derivations or detects properties) and R (which provides content). Its role in guiding inference and evaluation based on these criteria necessitates its integration within the Trialectic system for achieving states like justified belief."

Objection Finalized – The Risk of Trivializing Necessity

The critic might offer a final challenge, suggesting that by defining MPAASNI such that it requires the integrated functioning of R, I, and P, and then defining the Trialectics as precisely that integrated functioning, the thesis risks becoming true by definition, potentially lacking substantive content.

The final objection states: "You have defined MPAASNI in Section 4 such that its 'meaningful' and 'integrated' nature explicitly requires the coordinated operation of R, I, and P functions. You have defined the Trialectics in Section 3 as the name for this coordinated, mutually conditioning operation. In Section 5, you then argue that the Trialectics is necessary for MPAASNI because the definition of MPAASNI requires the functional integration that the Trialectics describes. Doesn't this structure make the necessity claim analytically true or true by stipulation based on your definitions? If 'MPAASNI' just *means* 'state produced by Trialectics', then of course Trialectics is necessary for MPAASNI. Have you demonstrated a substantive functional necessity, or merely unpacked the consequences of your chosen definitions?"

Response – Necessity Derived from Functional Analysis of the Phenomenon

The response must clarify that the necessity is not merely stipulative but arises from a functional decomposition of the *phenomenon itself* (MPAASNI), arguing that the definitions accurately capture the inherent functional requirements of that phenomenon.

The framework concludes: "This is a crucial methodological point concerning the nature of philosophical analysis. The claim is indeed that the necessity follows from the definitions, but the definitions themselves are not arbitrary stipulations. They are presented as the result of a

functional analysis of the target phenomenon: meaningful propositional attitudes about something not itself, as we understand them pre-theoretically and as they function in our cognitive lives and philosophical discourse (belief, desire, etc.).

The process was:

1. Identify the phenomenon (MPAASNI – states like belief, desire).
2. Analyze the functionally necessary components required to constitute *any* state exhibiting the core features of this phenomenon (directed content -> requires R function; systematic processing -> requires I function; internal stance/weighting -> requires P function). This decomposition is based on analyzing what *must* be involved for something to *be* a belief or desire, etc.
3. Analyze the *relationship* required between these functional components for them to constitute a unified, integrated state that matches the phenomenon (requiring mutual conditioning for functional significance and coherence -> Trialectics).
4. The necessity claim then connects the identified necessary functional architecture (Trialectics) to the phenomenon (MPAASNI) it enables.

The definitions are intended to capture the *functional essence* of MPAASNI. If the functional analysis presented in Section 4 – identifying directed content, systematic processability, internal stance, and their necessary integration as constitutive of MPAASNI – is accurate, then the necessity of the Trialectics (as the description of that integration) follows rigorously. The argument's substance lies in the claim that *this specific functional decomposition and the requirement for integration accurately capture the necessary conditions for the phenomenon itself*.

The thesis is therefore substantive, not trivial, because it makes a claim about the underlying functional structure required for a recognizable and significant class of cognitive states. It asserts that any system, regardless of its implementation, exhibiting states with the functional profile of meaningful belief, desire, etc., must instantiate this specific pattern of Trialectic interdependence. The necessity is conceptual, but it is grounded in an analysis of the functional requirements of the target phenomenon. The definitions aim to articulate these requirements precisely, and the necessity follows from that articulation. The debate, then, should focus on whether the functional analysis of MPAASNI (leading to the definitions) is indeed accurate and captures the necessary features of such states."

7.3. Examination of Objection 3: The Problem of Implementation and Empirical Underspecification

Having addressed potential objections concerning the necessity of mutual conditioning (Trialectics vs. simpler interaction) and the distinctness of Preference as a component, a third line of critique may arise regarding the framework's relationship to empirical implementation and its apparent abstractness. Critics might argue that the Trialectics, defined purely at a functional level, is too abstract to be empirically testable or to provide concrete insights into actual cognitive mechanisms. They might suggest that its high level of generality makes it compatible with potentially vastly different underlying physical or computational systems, thus lacking predictive power or failing to constrain possible implementations in a meaningful way. This objection questions the framework's explanatory value beyond purely conceptual clarification.

The Objection – Abstraction and Lack of Empirical Constraint

The objection proceeds: "The proposed Trialectic framework operates at a highly abstract functional level, defining Reference, Inference, and Preference based on their roles in constituting Meaningful Propositional Attitudes About Something Not Itself (MPAASNI). The core thesis posits a necessary functional architecture involving mutual conditioning. However, this description appears entirely divorced from any specific implementation details. You state the framework applies regardless of whether the system is biological, artificial, or collective, and regardless of the specific algorithms or physical structures used.

This level of abstraction raises concerns about the framework's empirical content and explanatory power. If the Trialectics is compatible with virtually *any* sufficiently complex underlying mechanism that manages to produce states meeting the functional definition of MPAASNI, then what does the framework actually explain or predict about those mechanisms? Does it provide any constraints on cognitive science, neuroscience, or artificial intelligence research seeking to understand or build systems capable of belief, desire, and other attitudes?

Furthermore, could the very notion of neatly separating functions into R, I, and P, and describing their 'mutual conditioning', be an artifact of our high-level analysis rather than reflecting distinct, interacting functional realities at the implementation level? Neural systems, for instance, often exhibit highly distributed and overlapping processing, where functions related to representation (R), computation (I), and valuation/motivation (P) might be deeply intertwined in ways not easily captured by a model of three distinct capacities conditioning each other.

The objection, therefore, is that the Trialectic framework, while perhaps internally consistent as a conceptual analysis, is empirically underspecified. It identifies abstract functional necessities but fails to connect adequately with the concrete mechanisms that realize these functions. Its strength (generality across implementations) becomes its weakness (lack of specific explanatory or predictive power regarding those implementations). Does the framework offer more than a redescription of the phenomenon (MPAASNI requires integrated content, processing, and stance) using the novel terminology of "Trialectics"?"

Response – Functional Level Analysis and Constraints on Architecture

The response must clarify the nature and value of functional-level analysis in philosophy and cognitive science, arguing that the Trialectics framework, while abstract, does provide non-trivial constraints on possible implementations and offers a valuable conceptual structure for understanding the necessary organization of functions, even if it doesn't specify algorithms or neural circuits.

The framework responds: "This objection highlights the distinction between functional-level description and implementation-level specification, a distinction crucial in cognitive science and philosophy of mind (e.g., David Marr's levels of analysis in *Vision*, 1982). The Trialectic framework deliberately operates at the functional or computational level (in Marr's sense), describing *what* functions must be performed (R, I, P) and *how* they must be organized (mutual conditioning via Trialectics) for a particular capability (MPAASNI) to be possible. This level of analysis is distinct from, but complementary to, the analysis of algorithms and physical implementation.

The value of this functional-level analysis is not undermined by its abstraction; rather, its value lies precisely in identifying necessary organizational principles that must hold *regardless* of the specific implementation details. The claim that the Trialectics is necessary for MPAASNI is a claim about the required *functional architecture*. It asserts that any system, no matter how implemented, must instantiate this pattern of mutual conditioning among its R-related, I-related, and P-related functions if it is to exhibit meaningful propositional attitudes.

This *does* impose constraints, albeit abstract ones, on possible implementations:

- **Constraint on Modularity:** It rules out architectures where R, I, and P functions are realized by strictly modular, non-mutually-conditioned components (as argued against Objection 1). Any successful implementation must allow for the dynamic, reciprocal influence between processes handling content-direction, content-processing, and content-evaluation/stance.
- **Constraint on Integration:** It implies that the mechanisms realizing R, I, and P cannot be entirely segregated. There must be pathways for information flow and mutual influence that embody the conditioning relations described. Whether this is achieved through overlapping neural circuitry, shared computational resources in an AI, or specific communication protocols in a collective system, the functional requirement for this interconnectedness remains.
- **Constraint on Function:** It specifies *what kinds* of interactions must be supported. For example, the mechanisms supporting Preference must be sensitive to inputs related to referential content and inferential processing, and the mechanisms supporting Inference must be sensitive to guiding signals related to Preference.

Regarding the concern about distributed processing in neural systems: The Trialectics describes the interrelation of *functions*, not necessarily anatomically distinct brain regions or computationally distinct algorithms. It is entirely possible for these functions (establishing reference, performing inference, applying preference) to be implemented in a highly distributed and overlapping manner within a neural network, provided that the *overall dynamic* of the network exhibits the required pattern of mutual conditioning. For instance, neural activity related to representing an object (R) might simultaneously be involved in computing its likely properties (I) and evaluating its salience or value (P), with feedback loops ensuring that the representation, computation, and valuation mutually influence each other. The Trialectics describes this functional pattern of interaction, not the specific location or nature of the realizing hardware.

Therefore, the framework is not empirically empty. It provides a high-level blueprint of the necessary functional organization. While it doesn't predict specific neural firing patterns or algorithms, it predicts that any system successfully implementing MPAASNI will exhibit functional interdependencies corresponding to the Trialectics. This can guide empirical research by suggesting what kinds of functional interactions to look for (e.g., how valuation systems influence representational areas and reasoning circuits, and vice versa) and provides a conceptual framework for interpreting findings. It also constrains AI design by indicating that achieving genuine belief and desire (as opposed to mere mimicry) requires building systems with this deep functional integration, not just separate modules for representation, logic, and goal-setting.

The Trialectics offers more than redescription; it offers a specific hypothesis about the *necessary structure of functional integration* underlying a crucial class of cognitive states. Its

abstractness allows it to capture a potentially universal principle of complex cognitive organization, while still providing meaningful constraints on the types of architectures capable of supporting such states."

Objection Refined – The Difficulty of Operationalizing and Testing Mutual Conditioning

The critic might accept the distinction between functional and implementation levels but press further on the practical difficulty of verifying the presence of the Trialectics, particularly the notion of "mutual conditioning," even at a functional level.

The refinement proceeds: "Granting that the framework operates at a functional level and posits architectural constraints, how can we reliably determine whether a given system actually instantiates the specific 'mutual conditioning' described by the Trialectics? The definitions of the conditioning pathways (e.g., P guiding R, I refining R, I informing P) are complex and interwoven. Distinguishing genuine mutual conditioning from complex but ultimately feedforward or centrally mediated interactions (as discussed in Objection 1, Move 3) seems extremely difficult in practice, whether analyzing a biological system or an AI.

If the core concept of 'mutual conditioning' that defines the Trialectics is hard to operationalize – hard to specify clear criteria for identifying its presence or absence based on observable system behavior or internal processing (even at an abstract functional level) – then the necessity claim, while perhaps conceptually coherent, becomes practically unverifiable. We might be unable to definitively say whether a system exhibits the Trialectics or merely a sophisticated approximation. Doesn't this difficulty in operationalizing the central mechanism weaken the claim that it represents a well-defined necessary condition? How do we move from the abstract description of mutual conditioning to criteria that would allow us to recognize its instantiation?"

Response – Functional Dependencies and Counterfactual Analysis

The response must address the operationalization challenge by suggesting ways to identify the presence of mutual conditioning through analyzing functional dependencies and potentially using counterfactual reasoning, even at an abstract level.

The framework responds: "The challenge of operationalizing complex functional concepts like 'mutual conditioning' is significant but not insurmountable, even within a conceptual or functional analysis. Identifying the Trialectics involves looking for specific patterns of functional dependency, which can be conceptually probed.

The presence of mutual conditioning can be identified by examining *functional dependencies* and *counterfactual scenarios*:

- **Dependency of R on I/P:** Does altering the system's active inference patterns or preference states systematically change its referential behavior (e.g., how it categorizes objects, which terms it applies, its precision of reference)? If changes in I or P reliably lead to specific changes in R's operation, this indicates conditioning ($I/P \rightarrow R$).
- **Dependency of I on R/P:** Does altering the specific referential content being processed or the active preference state systematically change the inferential pathways activated,

the conclusions drawn, or the speed/efficiency of inference? If changes in R or P reliably lead to specific changes in I's operation, this indicates conditioning ($R/P \rightarrow I$).

- **Dependency of P on R/I:** Does altering the referential content being considered or the outcome of inferential processing systematically change the system's preference state (e.g., its evaluation of an object, its commitment to a belief, its desire for an outcome)? If changes in R or I reliably lead to specific changes in P's state, this indicates conditioning ($R/I \rightarrow P$).

Crucially, the *mutual* aspect requires demonstrating dependencies in *all* these directions. The "conditioning" implies more than just receiving input; it implies that the *internal state or mode of operation* of the conditioned component is altered by the state of the conditioning component.

Counterfactual Analysis: We can conceptually analyze hypothetical scenarios. If system S exhibits MPAASNI, we ask: What would happen to its capacity for MPAASNI if we hypothetically severed the functional link allowing P to guide I? According to the framework, meaningful attitude formation or revision requiring preference-guided inference should break down. What if we severed the link allowing I to inform P? Preference states might become less responsive to evidence or logical consequences. What if R was not guided by P? Referential focus might become arbitrary or irrelevant. If these hypothetical functional lesions impair the system's ability to form or maintain integrated MPAASNI in the ways predicted by the Trialectic model, it supports the claim that these specific interdependencies are necessary.

While applying these analyses rigorously to complex real-world systems is empirically challenging, the conceptual criteria for identifying mutual conditioning *at the functional level* are specifiable. They involve looking for systematic, reciprocal dependencies between the operations associated with reference, inference, and preference. The Trialectics is operationalized not by finding specific neural correlates, but by identifying this pattern of functional interdependence. The difficulty lies in the complexity of the system and our access to its internal states, not necessarily in the conceptual coherence or identifiability of the Trialectic pattern itself. The framework provides the conceptual tools to recognize the pattern if the functional dependencies can be mapped. Therefore, the claim remains a well-defined assertion about a necessary functional architecture, even if verifying its instantiation empirically is complex."

Objection Further Refined – Gradations of Conditioning and the Necessity Threshold

The critic might acknowledge the possibility of identifying functional dependencies but argue that conditioning comes in degrees. They might question whether the necessity claim requires a *maximal* or *pervasive* form of mutual conditioning, or if weaker forms might suffice for *some* states qualifying as MPAASNI, suggesting the framework's sharp necessity claim is too rigid.

The refinement proceeds: "Your response suggests identifying mutual conditioning through functional dependencies. However, conditioning is likely not an all-or-nothing affair. The degree to which Preference guides Inference, or Inference refines Reference, might vary depending on the specific attitude, the context, or the system's developmental stage. Could there be a spectrum of functional integration?"

Perhaps rudimentary MPAASNI require only weak or partial mutual conditioning, while highly complex, rational, or self-reflective attitudes require the full, pervasive Trialectic interrelation you describe. For instance, a basic desire might involve strong $P \rightarrow I/R$ guidance but very weak $I/R \rightarrow P$ feedback. A deeply entrenched, systematically justified belief might exhibit strong conditioning in all directions.

If functional integration admits of degrees, and if the threshold for a state qualifying as 'meaningful' MPAASNI is potentially fuzzy or context-dependent, then claiming the *specific, fully described Trialectic architecture* is necessary for *all* MPAASNI seems problematic. It imposes a potentially artificial uniformity on a potentially diverse range of cognitive states. Shouldn't the necessity claim be qualified, acknowledging that perhaps only *partial* or *developing* forms of the Trialectics are necessary for the *emergence* of MPAASNI, while the fully elaborated version is necessary only for their most sophisticated forms? This would make the framework more nuanced but might weaken the strong necessity claim tied to the specific, fully described Trialectic structure."

Response – Trialectics as Necessary Core Structure, Admitting Variation in Strength/Focus

The response must address the issue of gradations by arguing that while the *strength* or *specific manifestation* of the conditioning pathways might vary, the underlying *structural pattern* of mutual conditioning itself remains necessary for any state qualifying as MPAASNI, ensuring the required functional integration, however simple or complex the attitude.

The framework concludes: "This objection concerning gradations is pertinent and requires careful consideration of the relationship between the core Trialectic structure and its varied manifestations in specific attitudes. The framework can accommodate variation in the strength or focus of the conditioning pathways without abandoning the necessity of the overall mutual conditioning pattern for *any* state qualifying as MPAASNI.

The key lies in the *functional integration* requirement (4.6d) derived from the definition of MPAASNI. Even for a seemingly simple attitude like a basic belief or desire to be *meaningful* and *about something not itself*, it must possess integrated directed content, systematic processability (however rudimentary), and an internal stance/weighting. This integration, it is argued, necessitates that the R, I, and P functions are not operating in isolation but are conditioning each other.

- **Minimum Threshold for Integration:** Consider the minimal requirements. For content (R) to be functionally integrated with stance (P), P must be directed *at* that specific R-content ($R \rightarrow P$), and the P-stance must potentially influence how R is used or focused ($P \rightarrow R$, e.g., attending to the believed object). For content (R) to be integrated with processing (I), R must provide input to I ($R \rightarrow I$), and I's processing must potentially refine or structure R ($I \rightarrow R$, e.g., classifying the referent). For processing (I) to be integrated with stance (P), I must potentially inform P ($I \rightarrow P$, e.g., evidence influencing belief strength), and P must potentially guide I ($P \rightarrow I$, e.g., goal directing practical inference or applying basic consistency checks). Even at a minimal level, these mutual dependencies seem required for the state to be a unified attitude rather than disconnected components. The absence of *any* of these conditioning links would lead to the functional disintegration argued against previously. Therefore, the *pattern* of

mutual conditioning – the Trialectics itself – appears necessary even for rudimentary MPAASNI.

- **Variation within the Structure:** The framework allows for significant variation in the *strength*, *bandwidth*, or *specific content* of the conditioning pathways in different attitude configurations.
 - A simple perceptual belief might involve strong $R \rightarrow I$ and $P \rightarrow I$ (based on sensory preference) but perhaps weaker $I \rightarrow P$ conditioning initially.
 - A complex theoretical belief might involve very strong $I \rightarrow P$ (justification based on inference) and strong $P \rightarrow I$ (guidance by epistemic values like coherence).
 - A basic desire might involve strong $P \rightarrow I/R$ but initially weak $I/R \rightarrow P$ conditioning, which might strengthen as the system learns about the consequences or achievability of the desire through inference and reference.
 - Sophistication arises not from adding new fundamental components, but from the richness, complexity, and potentially recursive application of the conditioning pathways within the Trialectic structure. Higher-order thought, for instance, involves the Trialectics operating on contents that refer to states produced by the Trialectics itself.
- **Defining "Meaningful":** The threshold for "meaningful" implies sufficient integration for the state to play its characteristic functional role. A state based on such weak or partial conditioning that it cannot reliably guide action (if a desire), inform inference (if a belief), or be subject to minimal consistency checks might fall below the threshold of being a *meaningful* propositional attitude *within the system's economy*. The Trialectics describes the necessary structure for crossing that functional threshold.

Therefore, the response maintains that the *structural pattern of mutual conditioning (Trialectics) is necessary for all MPAASNI*, providing the minimal required functional integration. The undeniable variations in complexity, strength, and specific dynamics among different attitude types are accommodated *within* this necessary structure as different configurations or states of the Trialectic system, particularly involving variations in the P-component and the weighting of the conditioning pathways. The necessity claim pertains to the fundamental architecture enabling the integration, not to a specific, uniform strength of interaction in all cases. The core structure of mutual conditioning remains indispensable."

8. Conclusion

8.1. Summary of Findings

This paper embarked on a conceptual analysis, operating under strict descriptive and methodological constraints, to investigate the functional architecture necessary for the possibility of complex cognitive states known as Meaningful Propositional Attitudes About Something Not Itself (MPAASNI). The analysis proceeded through several incremental stages:

First, we identified and provided functional definitions for three core capacities posited as necessary constituents of any system capable of exhibiting MPAASNI:

- **Reference (R):** The capacity for directed content, enabling internal states to be *about* entities, properties, or states of affairs distinct from the system itself.
- **Inference (I):** The capacity for systematic processing and relation of contents, enabling reasoning, derivation of conclusions, and the structuring of propositional networks.

- **Preference (P):** The capacity for internal stance, weighting, orientation, or selection regarding contents or processes, providing the specific mode of an attitude and its functional significance within the system.

Second, we articulated the concept of the *Trialectics*, defined not as a component but as the specific functional *mode of interrelation* characterized by *systematic and constitutive mutual conditioning* among R, I, and P. We detailed the various pathways through which R conditions and is conditioned by I and P; I conditions and is conditioned by R and P; and P conditions and is conditioned by R and I. This Trialectic interrelation was presented as a necessary description of the functional architecture required for the deep integration of directed content, systematic processing, and internal stance.

Third, we provided a precise functional definition of the target phenomenon, *Meaningful Propositional Attitudes About Something Not Itself (MPAASNI)*. This definition highlighted four necessary features: the presence of propositional content, directedness towards something not itself, functional significance (meaningfulness), and crucially, the *functional integration* of these aspects into a unified state.

Fourth, we presented the central argument demonstrating the *necessity* of the Trialectics for the possibility of MPAASNI. This argument proceeded by showing that the functional integration required by the very definition of MPAASNI logically entails the presence of the mutual conditioning among R, I, and P described by the Trialectics. Analysis of the contrapositive (absence of Trialectics) confirmed that without this specific mode of interrelation, the necessary functional integration cannot occur, rendering MPAASNI impossible within the framework's terms.

Finally, we explored the scope of this necessity, finding it applicable to any system exhibiting MPAASNI, regardless of substrate or whether the attitude's object is tangible or non-tangible. We accounted for the diversity of particular attitude types (belief, desire, hope, fear, etc.) as different configurations or patterns of activation within the necessary Trialectic architecture, often distinguished by the specific orientation of the Preference component. We also identified derivative implications, suggesting that other complex phenomena dependent on MPAASNI (like rational agency, complex communication, systematic knowledge acquisition) are also fundamentally impossible without the underlying Trialectic functioning. Through dialectical engagement with potential objections, the necessity of the distinct Preference component and the specific mutual conditioning of the Trialectics (as opposed to simpler interactions or external modulation) for the functional integration required by MPAASNI was defended.

8.2. Restatement of Thesis

The core finding of this conceptual investigation, derived strictly from the descriptive analysis of functional requirements, is encapsulated in the central thesis:

"The Trialectic Functioning of a System—defined as the systematic and constitutive mutual conditioning among its capacities for Reference, Inference, and Preference—is a Necessary Condition for the Possibility of that System Having Meaningful Propositional Attitudes About Something Not Itself."

8.3. Philosophical Significance

The philosophical significance of this thesis, should its conceptual grounding prove robust under further scrutiny, lies in several areas:

- **Foundation for Intentionality and Meaning:** It proposes a specific functional architecture as the basis for intentionality and meaningfulness in complex cognitive states, suggesting that directedness towards the world and functional significance within a system arise not just from representation or computation alone, but from the deeply integrated interplay of representation (R), computation (I), and internal orientation/value (P).
- **Understanding Cognitive Integration:** It highlights the necessity of functional integration, specifically mutual conditioning, for sophisticated cognition. It challenges models that posit strictly modular capacities for content, reasoning, and motivation, suggesting instead a fundamental entanglement at the functional level is required for states like belief and desire.
- **A Unified Framework:** The Trialectics offers a potentially unifying framework for analyzing diverse propositional attitudes (belief, desire, hope, fear, intention) by treating them as different configurations within the same necessary functional architecture, distinguished primarily by the role and orientation of the Preference component.
- **Constraints on Cognitive Modeling:** As a claim about necessary functional architecture, it provides abstract but potentially useful constraints for cognitive science and artificial intelligence. Building systems capable of genuine MPAASNI might require designing architectures that explicitly implement or allow for the emergence of this mutual conditioning among R, I, and P functions.
- **Reframing Epistemology and Rationality:** By embedding justification and knowledge within the context of the Trialectics, the framework suggests that epistemic standing depends not just on inference or evidence, but on the well-configured interplay of reference, inference, and *epistemic preferences*. Rationality itself might be reconceived as a property of the harmonious functioning of this integrated system.

8.4. Directions for Future Research

This paper lays conceptual groundwork; significant further work is required to develop, refine, and test the Trialectic framework:

- **Formal Modeling:** Exploring the possibility of developing formal models (perhaps drawing from systems theory, dynamic systems theory, or logic) that capture the mutual conditioning dynamics of the Trialectics more precisely.
- **Detailed Analysis of Specific Attitudes:** Applying the framework in greater detail to analyze the specific Trialectic configurations constituting less-explored attitudes (e.g., trust, regret, imagination, intention).
- **Application to Specific Philosophical Problems:** Investigating whether the Trialectics framework can shed new light on long-standing philosophical problems, such as the nature of consciousness, the frame problem in AI, the semantics of attitude reports, or the grounding of normativity.
- **Comparative Analysis:** Systematically comparing the Trialectic framework with existing models in philosophy of mind and cognitive science (e.g., specific versions of

RTM, predictive processing models, integrated information theory) to highlight similarities, differences, and potential advantages or disadvantages.

- **Exploring Development and Learning:** Investigating how the Trialectic functional architecture might develop or be learned within a system, and how the conditioning pathways might be modified through experience.
- **Investigating Pathologies:** Analyzing whether certain cognitive deficits or psychopathologies could be understood as specific disruptions or imbalances within the Trialectic system.

In conclusion, by focusing on the functional requirements for meaningful propositional attitudes directed beyond the system itself, this analysis has derived the Trialectics – the mutual conditioning of Reference, Inference, and Preference – as a necessary underlying architecture. While abstract, this framework offers a potentially powerful lens for understanding the deep functional integration required for sophisticated cognition and agency, inviting further exploration and critical engagement.

Glossary

- **About Something Not Itself (Directedness/Intentionality):** A defining characteristic of certain cognitive states (specifically MPAASNI within this framework) signifying that their content is directed towards entities, properties, relations, or states of affairs that are distinct from the cognitive state itself and often external to the system holding the state. This requires a functional link between the internal representation and the external or distinct target, a function fulfilled by Reference (R).
- **Belief:** A particular type of Meaningful Propositional Attitude About Something Not Itself (MPAASNI) characterized by an internal stance (Preference) of acceptance or affirmation regarding the truth or factuality of its propositional content. Within the Trialectics, the Preference component in belief is oriented towards truth-tracking and guides Inference and Reference accordingly.
- **Cognitive System:** An entity (e.g., individual agent, potentially an artificial intelligence or a structured collective) possessing capacities for processing information, forming internal states, and potentially interacting with an environment. Within this framework, the focus is on systems capable of exhibiting Reference, Inference, and Preference, and potentially MPAASNI.
- **Conceptual Analysis:** The philosophical method employed throughout the development of this framework, involving the clarification of concepts, the identification of necessary functional requirements, and the examination of logical relationships between defined terms and phenomena, without direct reliance on empirical experimentation.
- **Content (Propositional):** The structured, truth-evaluable information that forms the object of a Propositional Attitude. It is composed of constituent elements (concepts representing objects, properties, relations) linked via Reference and organized in a specific structure, often processed or related via Inference.
- **Descriptive Methodology:** The approach taken in defining terms and analyzing relationships, focusing on characterizing *what* functions are performed and *how* components interrelate, strictly avoiding normative judgments (e.g., 'good', 'rational', 'correct') or teleological claims (claims about inherent goals or purposes), except where 'goal' is used descriptively as part of the Preference function (e.g., desire aims at realization).
- **Desire:** A particular type of Meaningful Propositional Attitude About Something Not Itself (MPAASNI) characterized by an internal stance (Preference) oriented towards the realization or bringing about of the state of affairs described by its propositional content. Within the Trialectics, the Preference component in desire motivates goal-directed Inference and action-planning concerning the referentially constituted content.
- **Epistemology:** The branch of philosophy concerned with the nature, origin, scope, and justification of knowledge and belief. The Trialectic framework has implications for epistemology by proposing a necessary functional architecture for states like justified belief and knowledge.
- **Fear:** A particular type of Meaningful Propositional Attitude About Something Not Itself (MPAASNI) characterized by an internal stance (Preference) of apprehension oriented towards the non-realization or avoidance of a negatively evaluated potential state of affairs described by its propositional content. Within the Trialectics, the Preference component guides Inference towards risk assessment and avoidance.
- **Functional Architecture:** An abstract description of a system based on the functions performed by its components and the pattern of interactions and dependencies among them, independent of the specific physical or algorithmic implementation realizing those functions. The Trialectics is proposed as a necessary functional architecture for MPAASNI.

- **Functional Definition:** A definition of a concept (like Reference, Inference, Preference) based on the specific role or function it performs within a larger system or process, rather than its intrinsic nature or physical realization.
- **Functional Integration:** The state within a system where distinct functional capacities (specifically R, I, and P in this context) operate in a coordinated and interconnected manner, such that their operations are mutually influential and contribute jointly to constituting a unified cognitive state (like MPAASNI). This integration is described as requiring the mutual conditioning defined by the Trialectics.
- **Functional Significance (related to "Meaningful"):** The property of a cognitive state (like an MPAASNI) playing a discernible and systematic role within the system's overall cognitive or practical economy, influencing other states, processes, or potential interactions with the environment. This significance arises from the integrated operation of R, I, and P within the Trialectics.
- **Hope:** A particular type of Meaningful Propositional Attitude About Something Not Itself (MPAASNI) characterized by a complex internal stance (Preference) involving desire for the realization of a state of affairs combined with inferential assessment of its possibility, typically under conditions of uncertainty.
- **Inference (I):** Functionally defined as the capacity of a system to process and relate contents (typically referentially constituted propositional contents), deriving new contents from existing ones or establishing structural relationships between them according to systematic patterns or rules. Its necessary functional role is to provide systematic processing and relation of contents.
- **Internal Stance / Weighting / Orientation / Selection:** The core function attributed to Preference (P). It refers to the system's internal disposition towards a propositional content or state of affairs, determining the *mode* of the attitude (belief, desire, etc.) and its functional significance or weighting within the system's economy.
- **Meaningful Propositional Attitudes About Something Not Itself (MPAASNI):** The central class of complex cognitive states analyzed. Defined functionally as states possessing: (a) Directed Propositional Content (via R), (b) Systematic Processability (via I), (c) Specific Internal Stance and Functional Weighting (via P), and crucially (d) Functional Integration of these components, necessitating the Trialectic functioning. Examples include belief, desire, hope, fear, intention, etc., directed towards something distinct from the attitude.
- **Metaphysics:** The branch of philosophy concerned with the fundamental nature of reality, existence, being, and the world. The discussion of the "metaphysical status" of the Trialectics pertains to what kind of entity, process, relation, or property this functional architecture represents within a fundamental ontology.
- **Mutual Conditioning:** The defining characteristic of the Trialectics. It describes the state where the form, state, or operation of each component (R, I, P) is systematically dependent upon, constrained by, and influenced by the concurrent form, state, or operation of the other two components within the system, creating a dynamic, interdependent functional loop.
- **Necessary Condition:** A condition that *must* be present for a particular outcome or phenomenon to occur. The central thesis claims the Trialectics is a necessary condition for the *possibility* of MPAASNI – if a system exhibits MPAASNI, it *must* possess the Trialectic functional architecture.
- **Non-tangible Domains:** Realms of entities, properties, or states of affairs that are not physical or directly perceivable, such as abstract objects (numbers, concepts), possibilities, fictional entities, logical relations, or other mental states. The Trialectics is argued to be necessary for MPAASNI directed towards these domains as well.

- **Preference (P):** Functionally defined as the capacity of a system to exhibit an internal stance, weighting, orientation, or selection regarding contents, states of affairs, or processing pathways. Its necessary functional role is to provide the specific mode of an attitude and its functional significance or meaningfulness through internal weighting and guidance.
- **Proposition:** The abstract, structured, truth-evaluable content that serves as the object of a Propositional Attitude. It is composed of constituent elements linked via Reference.
- **Propositional Attitude:** A type of cognitive state characterized by an agent or system taking a specific stance or "attitude" (e.g., belief, desire, fear) towards a proposition. MPAASNI is a specific subclass of these.
- **Rational Agency:** The capacity of an agent to act for reasons, typically understood as involving acting on the basis of beliefs and desires (MPAASNI) through processes of practical inference. Argued to be impossible without the underlying Trialectics necessary for MPAASNI.
- **Reference (R):** Functionally defined as the capacity of a system to establish and utilize connections between its internal states or representational elements and specific entities, properties, relations, or states of affairs. Its necessary functional role is to provide directed content, enabling states to be *about* something.
- **Systematic Processing:** The function attributed to Inference (I), referring to the capacity to manipulate, relate, and derive contents according to regular patterns or rules, enabling reasoning and the structuring of knowledge.
- **Tangible Reality:** The realm of physical objects, events, properties, and relations existing in space and time and perceivable through the senses. The Trialectics is necessary for forming MPAASNI *about* this realm.
- **Trialectics:** The central concept of the framework, defined as the specific functional mode of operation within a system characterized by systematic and constitutive mutual conditioning among its capacities for Reference (R), Inference (I), and Preference (P). It represents the necessary functional architecture for the possibility of MPAASNI.

References

1. Audi, Robert. *Epistemology: A Contemporary Introduction to the Theory of Knowledge*. 3rd ed., Routledge, 2010.
2. Block, Ned, editor. *Readings in Philosophy of Psychology, Vol. 1*. Harvard University Press, 1980.
3. Brentano, Franz. *Psychology from an Empirical Standpoint* (Psychologie vom empirischen Standpunkte). 1874. English translation by Antos C. Rancurello, D. B. Terrell, and Linda L. McAlister, Routledge, 1973.
4. Dennett, Daniel C. *The Intentional Stance*. MIT Press, 1987.
5. Fodor, Jerry A. *The Language of Thought*. Harvard University Press, 1975.
6. Frege, Gottlob. "On Sense and Reference" ("Über Sinn und Bedeutung"). 1892. Translated in *Translations from the Philosophical Writings of Gottlob Frege*, edited by Peter Geach and Max Black, Blackwell, 1952.
7. Kim, Jaegwon. *Philosophy of Mind*. 3rd ed., Westview Press, 2010.
8. Kripke, Saul A. *Naming and Necessity*. Harvard University Press, 1980.
9. Lycan, William G. *Philosophy of Language: A Contemporary Introduction*. 3rd ed., Routledge, 2018.
10. Marr, David. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. W. H. Freeman, 1982.
11. Mill, John Stuart. *A System of Logic, Ratiocinative and Inductive*. 1843.
12. Peirce, Charles Sanders. *Collected Papers of Charles Sanders Peirce*. Edited by Charles Hartshorne, Paul Weiss, and Arthur W. Burks. Harvard University Press, 1931–1958.
13. Putnam, Hilary. "The Meaning of 'Meaning'". In *Mind, Language and Reality: Philosophical Papers, Vol. 2*. Cambridge University Press, 1975.
14. Resnik, Michael D. *Choices: An Introduction to Decision Theory*. University of Minnesota Press, 1987.
15. Russell, Bertrand. "On Denoting". *Mind*, vol. 14, no. 56, 1905, pp. 479–493.
16. Sainsbury, R. M. *Logical Forms: An Introduction to Philosophical Logic*. 2nd ed., Blackwell Publishing, 2001.
17. Searle, John R. *Intentionality: An Essay in the Philosophy of Mind*. Cambridge University Press, 1983.

